(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号
特開2002－208939
（P2002－208939A）
(43)公開日　平成14年７月26日(2002.7.26)

(51)Int.Cl.⁷	識別記号	ＦＩ	ﾃｰﾏｺｰﾄﾞ(参考)
Ｈ０４Ｌ　12/56	１００	Ｈ０４Ｌ　12/56	Ａ　５Ｋ０３０ １００Ａ

審査請求　未請求　請求項の数11　ＯＬ　外国語出願　（全 67 頁）

(21)出願番号　特願2001－357635(P2001－357635)

(22)出願日　平成13年11月22日(2001.11.22)

(31)優先権主張番号　２３２７８９６
(32)優先日　平成12年12月８日(2000.12.8)
(33)優先権主張国　カナダ（ＣＡ）

(71)出願人　501279833
アルカテル・カナダ・インコーポレイテッド
カナダ国、オンタリオ・ケー・２・ケー・２・イー・６、カナタ、マーチ・ロード・600
(72)発明者　マイク・リーブス
カナダ国、オンタリオ・ケー・１・エス・３・ワイ・６、オタワ、モンク・ストリート・16
(74)代理人　100062007
弁理士　川口　義雄　（外５名）

最終頁に続く

(54)【発明の名称】　ＡＴＭプラットフォームにおけるＭＰＬＳ実装

(57)【要約】
【課題】　通信ネットワーク内の第１のノードおよび第２のノードの間で接続パスを確立しようとする試行を計時する方法を提供すること。
【解決手段】　この方法は、それが存在する場合、以前の２回の、接続を確立しようとする試行間に前に経過した期間よりも長い期間が経過した後、接続パスを確立しようとする試行を開始する。

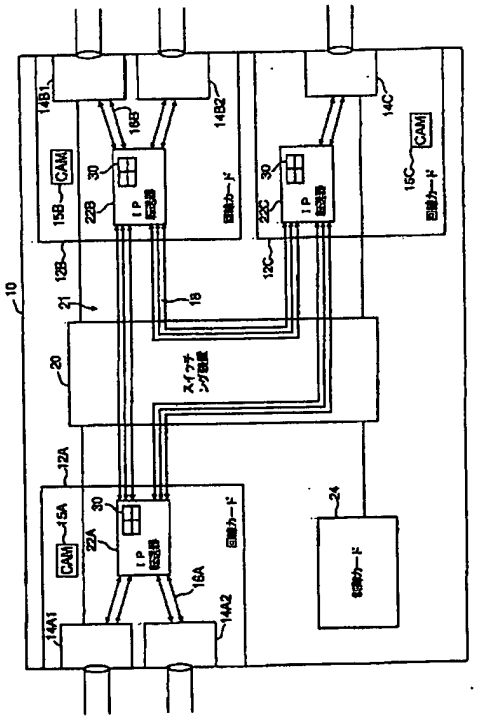

Figure 1

1

【特許請求の範囲】
【請求項１】　通信ネットワーク内の第１のノードと第２のノードの間で接続パスを確立しようとする試行を計時する方法であって、ある期間が経過した後に接続パスを確立しようとする試行を開始することを含み、前記期間は、前記接続を確立しようとする以前の２回の試行が存在した場合、その試行間で前に経過した別の期間よりも長い方法。
【請求項２】　前記期間が前記別の期間よりも固定時間値だけ長い請求項１に記載の方法。
【請求項３】　前記期間が最大時間値を超えない請求項１に記載の方法。
【請求項４】　前記接続パスがソフト永久ラベルスイッチパスである請求項１に記載の方法。
【請求項５】　前記固定時間値が１０秒である請求項２に記載の方法。
【請求項６】　通信ネットワーク内の接続を求める複数の要求に対する接続を確立しようとする試行を計時する方法であって、
一定の時間間隔の経過を追跡するタイマ構成を有することと、
接続を求める前記複数の要求に関連するレコードのリストを有することと、
前記リストから１つのレコードを選択することと、
前記１つのレコードに関連する接続を確立しようと試みることとを含み、かつ前記１つのレコードに関連する前記接続が確立された場合には、
前記レコードに成功したものとしてマークを付け、そうでなければ、前記一定の間隔で増大する連続する間隔で前記接続を確立しようと再試行することをさらに含む方法。
【請求項７】　前記リストから１つのレコードを選択することが、
レコードの前記リスト内に時間フィールドを有することと、
レコードの前記リスト内の各エントリごとに各前記一定の時間間隔で、
前記時間フィールド内の時間値を減分することと、
あるエントリに関して前記時間値がゼロである場合には、
前記エントリを前記１つのレコードとして選択することとを含む請求項６に記載の方法。
【請求項８】　連続する時間間隔で前記接続を確立しようと再試行するとき、前記連続する時間間隔は最大時間値を超えない請求項６に記載の方法。
【請求項９】　前記最大時間値が６０秒である請求項８に記載の方法。
【請求項１０】　２つのノード間で関連する少なくとも２つの通信リンクを有する前記２つのノードを含む通信ネットワーク内で、ラウンドロビンアルゴリズムを使用して前記２つのノード間で通知を行うため、前記少なくとも２つの通信リンクのうちの１つを選択する方法。
【請求項１１】　前記２つのノード間での通信には不十分なリソースを有する、またはそこで障害を有する前記少なくとも２つの通信リンクのうちのどの通信リンクも選択しないことをさらに含む請求項１０に記載の方法。

2

【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、デジタル通信システムに関し、より詳細には、非同期転送モード（ＡＴＭ）プラットフォーム上でマルチプロトコルラベルスイッチング（ＭＰＬＳ）を使用するネットワークノードの実装に関する。
【０００２】
【従来の技術】ＭＰＬＳは、インターネットプロトコル（ＩＰ）パケットを伝送するロバストな方式として、急速に業界でサポートを得ている。これは主に、ＭＰＬＳではパケットのパス上のすべてのルータまたはネットワークノードで、パケットの宛先ＩＰアドレスを検査する必要性がないからである。このため、ＭＰＬＳは、多くのネットワークの高速コア内で特に有用性を有する。高速コア内にＡＴＭスイッチングインフラストラクチャが存在する可能性が高いことを認識して、業界は、現在、ＡＴＭインフラストラクチャ上でＭＰＬＳを導入するための標準を策定中である。
【０００３】
【発明が解決しようとする課題】ほとんどの標準化の作業の性質と同様に、その焦点は、様々な当事者によって製造される機器の間での相互運用性を可能にするのに必要な機能上の特徴を定義することである。ただし、ＭＰＬＳ機能性を実装する際に多くの問題が生じる。これらには、（１）ＡＴＭインフラストラクチャ上での通知されたラベルスイッチパス（ＳＬＳＰ）の一般的な管理および保守、（２）ＳＬＳＰの確立障害時の手続き、（３）シグナリングリンクの管理、（４）シグナリングリンクまたは物理リンクに障害が起きたときの手続き、および（５）新しいシグナリングリンクの作成などの、ネットワークトポロジの様々な変更の下での手続きが含まれる。本発明は、これらの様々な問題に解決策を提供しようとするものである。
【０００４】
【課題を解決するための手段】本発明の一態様は、複数の中間ノードを介してイングレスノードからイグレスノードに至る接続パスが確立される、複数の相互接続されたノードを有する通信ネットワークを管理する方法を提供する。この方法は、接続パスをネットワーク全体にわたる固有識別子と関連付けるステップと、そのパス識別子をイングレスノード上に記憶し、パスがそこから出ていることを示すようにするステップと、そのパス識別子を各中間ノード上に記憶し、そのパスがそのような各中

間ノードを通過することを示すようにするステップと、そのパス識別子をイグレスノード上に記憶し、そこにそのパスが終端することを示すようにするステップとを含む。

【０００５】好ましくは、接続識別子を記憶するステップは、中間ノードを介するイングレスノードからイグレスノードに至る接続セットアップ要求の通知による接続パスの確立中に行われる。

【０００６】本発明の別の態様は、初回に失敗しているＳＬＳＰなどの接続パスを確立しようとする試行を計時する方法に関する。これは、ある期間が経過した後、接続パスを確立しようとする新たな試行を開始することによって実現され、前記期間は、前記接続を確立しようとする試行が以前にあった場合、その２回の試行間で以前に経過した期間よりも長い。

【０００７】本発明の別の態様は、通信ネットワークにおける、ＳＬＳＰなどの接続を求める複数の要求に対する接続を確立しようとする試行を計時する方法に関する。この方法は、定常時間間隔の経過を追うためのタイマ構成を提供するステップと、接続を求める複数の要求に関連するレコードのリストを提供するステップと、そのリストから１つのレコードを選択するステップと、その１つのレコードに関連する接続を確立しようと試行するステップと、その１つのレコードに関連する接続が確立された場合、その１つのレコードに成功したものとしてマークを付け、そうでなければ、定常間隔ずつ長くなる間隔で連続してその接続を確立しようと再試行するステップとを含む。

【０００８】その他の態様では、本発明は、前述した態様の様々な組み合わせおよびサブセットを提供する。

【０００９】本発明の前述した態様およびその他の態様は、以下に記す本発明の特定実施形態の説明、および本発明の原理を例としてのみ示す添付の図面からより明白となる。図面では、同じ要素には同じ参照番号が付いており、これらの参照番号は、同じ要素の特定の例示を識別するため、固有の英字サフィックスが付いている場合がある。

【００１０】

【発明の実施の形態】以下の説明およびそこでの実施形態は、本発明の原理の特定実施形態の１つまたは複数の例を示すものとして提供する。これらの例は、それらの原理を説明する目的で提供するものであり、それらの原理を限定するものではない。下記の説明では、本明細書および図面の全体にわたり、同じ要素には、それぞれ同じ参照番号を付けている。

【００１１】１．　ＡＴＭスイッチングの概要

図１は、例としての二重機能のＡＴＭスイッチおよびＩＰルータ１０（以降、「ノード」）のアーキテクチャを示すブロック図である。ノード１０は、物理インターフェース入力／出力ポート１４を有する回線カード１２などの、複数の入力／出力コントローラを含む。一般的に言って、回線カード１２は、ポート１４で着信ＡＴＭセルを受信する。標準化されたＡＴＭ通信プロトコルによれば、各ＡＴＭセルは、固定サイズのものであり、また仮想パス識別子（ＶＰＩ）および仮想チャネル識別子（ＶＣＩ）を組み込んで、セルを特定の仮想回路（ＶＣ）と関連付けることができるようにしている。受信したそのような各セルごとに、回線カード１２は、ＶＣ上の対応するルックアップテーブルまたはコンテントアドレッサブルメモリ（ＣＡＭ）１５を照会する。ＣＡＭ１５は、各セルごとに、発信ポートおよびイグレス回線カードに関する事前構成されたアドレス指定情報を提供する。これは、「イグレス接続インデックス」を使用して実現され、このインデックスは、セルが次のネットワークリンク上に進む際、そのセルに帰属させられるべき新しいＶＣ識別子を記憶するイグレス回線カードに関する事前構成されたメモリロケーションに対するポインタである。イングレス回線カードは、アドレス指定情報およびイグレス接続インデックスを各セルに付加して、このセルをスイッチング装置２０に送信し、このスイッチング装置は、このセルを適切なイグレス回線カードに物理的に再送する、またはコピーする。イグレス回線カードは、次に、事前構成されたＶＰＩ／ＶＣＩフィールドの置き換えを行い、セルをイグレスポートから送信する。この型のＡＴＭスイッチング機構のさらなる詳細は、ＰＣＴ公開番号ＷＯ９５／３０３１８で見ることができる。この公開番号はその全体が、参照により、本明細書に組み込まれる。

【００１２】また、ノード１０は、以下にずっと詳細に説明するとおり、経路指定機能およびシグナリング機能を含む様々なノード機能を制御および構成するための制御カード２４も備えている。回線カード１２は、スイッチング装置２０を介して、ポート１４で受信したデータをこの制御カード２４に送信することができる。

【００１３】各回線カードは、双方向のトラフィックフローをサポートする（すなわち、着信パケットおよび発信パケットを処理することができる）。ただし、説明では、以下の議論は、図１で左から右に流れるデータトラフィックに関して、回線カード１２Ａおよびポート１４Ａ１および１４Ａ２がイングレス処理を提供し、回線カード１２Ｂ、１２Ｃおよびポート１４Ｂ１、１４Ｂ２、１４Ｃ１、１４Ｃ２がイグレス処理を提供するものと想定する。

【００１４】２．　ＩＰ経路指定の概要

図示する実施形態のノードは、また、インターネットプロトコル（ＩＰ）などの階層的により高い通信レイヤと関連するデジタルデータの可変長パケットが、ＡＴＭトランスポートレイヤインフラストラクチャ上で搬送され得るようにもする。これは、各可変長パケットをトランスポートのために複数のＡＴＭセルにセグメント化する

ことにより、可能になる。したがって、あるＶＣは、ＩＰパケットを搬送することに専用にすることができ、他方、別のＶＣは、本来のＡＴＭ通信と排他的に関連しているようにすることができる。

【００１５】セルがイングレスポート１４Ａ１に着信すると、回線カード１２Ａは、ＣＡＭ１５Ａにアクセスして、前述したとおり、着信セルのＶＣに関するコンテキスト情報を入手する。このコンテキスト情報は、ＶＣを「サービスインターフェース」と関連付けることができる。これは、ネットワークを通るＡＡＬ５　ＡＴＭパスなどのリンクレイヤ（すなわち、「レイヤ２」）パスの端点である。各Ｉ／Ｏポート１４上にいくつかのサービスインターフェース（ＳＩ）が存在することが可能である。これらのサービスインターフェースは、同じ回線カード上のＩＰ転送器２２で「終端する」。これは、後述するとおり、ＩＰパケットを構成するＡＴＭセルがパケットに再組立てされ、その後に、ＩＰ転送手続き（ＡＴＭスイッチング手続きに対立するものとして）が行われる。

【００１６】ＩＰ転送の要点は、１つのＳＩで受信されたＩＰパケットが別のＳＩで送信されることである。図２に示すプロセス流れ図をさらに参照すると、ＩＰパケットに関する転送プロセスは、ノードを介して、２つの処理段階で隔てられた３つのトランスポート段階に論理的に分割することができる。

【００１７】矢印１６Ａによって概略で表す第１のトランスポート段階は、イングレスＳＩと関連するＡＴＭセルをイングレスポート１４Ａ１からイングレスＩＰ転送器２２Ａに搬送する。

【００１８】第２のトランスポート段階は、スイッチング装置２０を介して、イングレス転送器２２ＡからイグレスＩＰ転送器、例えば、転送器２２Ｂに、ＩＰパケットを搬送する。この第２のトランスポート段階は、「接続メッシュ」２１を介して実装される。この接続メッシュ内では、ＩＰ転送器２２の各対の間で、８つの内部接続、つまりトランスポートインターフェース（ＴＩ）１８がセットアップされる（３つのＴＩだけを示している）。これらのＴＩは、ＩＰパケットに対する異なるサービスレベルまたはサービスクラス（ＣＯＳ）を可能にするように提供されている。

【００１９】矢印１６Ｂによって概略で示す第３のトランスポート段階は、イグレスＩＰ転送器２２Ｂからイグレスポート、例えば、ポート１４Ｂ１、およびイグレスＳＩにＩＰパケットを搬送する。

【００２０】第１の処理段階は、イングレスＩＰ転送器２２Ａで行われ、そこで、イングレスＳＩと関連するＡＴＭセルがＩＰパケットに再組立てされる。これは、図２でステップ「Ａ」として示している。次に、ステップ「Ｂ」で、ＩＰ転送器２２Ａは、ネットワーク上の「次のホップ」のための適切なイグレスＳＩを決定するため、パケットの宛先ＩＰアドレスを検査する。この判定は、図３に概略で示すＩＰ転送テーブル３０（下記にさらに詳細に説明するとおり、ＩＰプロトコルから導出される）に基づく。テーブル３０の各レコードは、ＩＰアドレスフィールド３２および「イグレスインターフェースインデックス」フィールド３６を含む。パケットの宛先ＩＰアドレスをＩＰアドレスフィールド３２内でルックアップして、それに対する最長マッチ（すなわち、可能な限り遠くまでのパケットＩＰアドレスの宛先を解決するテーブルエントリ）を探し出す。対応するイグレスインターフェースインデックスは、実質的に、そのパケットに関するイグレス回線カード１２Ｂ、イグレスＩＰ転送器２２Ｂ、およびイグレスＳＩを特定する（より詳細には、図８Ｂに関連する議論を参照）。イグレスインターフェースインデックスは、ＩＰパケットに付加されている。

【００２１】さらに、ステップ「Ｃ」で、ＩＰ転送器２２Ａは、パケットによってカプセル化されたサービスクラス（ＣＯＳ）を検査する。部分的にはカプセル化されたＣＯＳおよび内部構成に基づいて、ＩＰ転送器２２Ａは、第２の段階のＴＩ１８のうちの１つを選択し、これが、所望のサービスクラスをもつイグレス転送器２２Ｂに到達する。スイッチング装置２０を通過するため、イングレスＩＰ転送器２２Ａは、ＩＰパケットをＡＴＭセル（ステップ「Ｄ」として概略で示す）に再セグメント化し、その宛先がイグレスＩＰ転送器２２Ｂであることを示すアドレス指定情報を各セルに付加する。

【００２２】第２のより小さい処理段階は、イグレスＩＰ転送器２２Ｂで行われ、そこで、パケットからイグレスインターフェースインデックスが抽出され、ステップ「Ｅ」で、イグレスＳＩと関連するカプセル化とマッチするように変更される。したがって、イグレスＳＩと関連するＶＰＩ／ＶＣＩが、そのパケットに付加される。次に、パケットは、イグレスＳＩ（「Ｇ」とラベル付けした）に対応する第３の段階のトランスポート１６Ｂを使用して送達される。このプロセスで、パケットは、再びＡＴＭセルにセグメント化され、このセルは、イグレスＳＩおよび／または出力ポート１４Ｂ１と関連するセル待ち行列内にバッファリングされる。また、待ち行列化が行われ、輻輳の起きる可能性のあるポイント（「Ｆ」とラベル付けした）も、第２の処理段階と第３のトランスポート段階の間、つまり、イグレスＩＰ転送モジュール２２Ｂとイグレスインターフェース（「Ｇ」とラベル付けした）の間で生じる。

【００２３】以上のことから、ＡＴＭプラットフォーム上でＩＰ転送機能性を実施することは、比較的複雑なプロセスであることが理解されよう。このプロセスは、ＩＰパケットが、（ａ）イングレスＩＰ転送器２２Ａで再組立てされることと、（ｂ）次に、スイッチング装置上でのトランスポートのためにセグメント化されること

と、（ｃ）イグレス転送器２２Ｂで再組立てされることと、（ｄ）次に、出力ポートからの送信のために再セグメント化されることを必要とする。さらに、イングレス転送器２２Ａで、トリビアルではないＩＰアドレスルックアップが行われなければならない。これらのステップは、各ネットワークノードで行われなければならず、したがって、終端間通信の待ち時間を増大させる。

【００２４】３．　ＭＰＬＳ概説

すべてのパケットに対してそれぞれ、前述の手続きを行わなければならないことを回避するため、ノード１０は、マルチプロトコルラベルスイッチング（ＭＰＬＳ）機能を提供する。従来のＩＰ転送では、各パケットの宛先アドレスに関する最長マッチであるアドレスプレフィックスがルータのテーブル内に存在する場合、２つのパケットは同一の「転送等価クラス」（ＦＥＣ）内にあるものと、通常、ルータはみなす。各ルータは、それぞれ独立にパケットを再検査し、パケットをＦＥＣに割り当てる。これとは対照的に、ＭＰＬＳでは、パケットがＭＰＬＳドメインに入った際に一度だけ、パケットをＦＥＣに割り当て、そのＦＥＣを表す「ラベル」をパケットに付加する。ＭＰＬＳがＡＴＭインフラストラクチャ上に導入される場合、ラベルは、特定のＶＣ識別子である。ＭＰＬＳドメイン内の後続のホップでは、ＩＰパケットは、もはや検査されない。代りに、ラベルが、次のホップを特定するテーブルへのインデックスおよび新しいラベルを提供する。したがって、ＭＰＬＳドメイン内の後続のホップでは、パケットを構成するＡＴＭセルは、従来のＡＴＭ技法を使用して交換することができる。そのようなパスは、当分野では、ラベルスイッチパス（ＬＳＰ）として知られ、ＬＳＰは、ネットワークオペレータによって手作業で、永久ラベルスイッチパス（ＰＳＬＰ）としてセットアップされることが可能である。あるいは、ネットワークオペレータからコマンドがあった際にネットワークがパスを自動的にセットアップする、ラベル配布プロトコル（ＬＤＰ）を使用することも可能である。そのようなパスは、通常、当分野では、ソフト永久ＬＳＰまたは通知されたＬＳＰ（ＳＬＳＰ）と呼ばれる。ＭＰＬＳに関するさらなる詳細は、以下のＭＰＬＳ標準案（すなわち、進行中）またはＭＰＬＳ提案で見ることができ、これらはそれぞれ、参照により、本明細書に組み込まれる。

［１］　Ｅ．Ｒｏｓｅｎ，　Ａ．Ｖｉｓｗａｎａｔｈａｎ，　Ｒ．Ｃａｌｌｏｎ，　Ｍｕｌｔｉｐｒｏｔｏｃｏｌ　Ｌａｂｅｌ　Ｓｗｉｔｃｈｉｎｇ　Ａｒｃｈｉｔｅｃｔｕｒｅ，　ｄｒａｆｔ　ｉｅｔｆ－ｍｐｌｓ－ａｒｃｈ－０６．ｔｘｔ．

［２］　Ｌ．Ａｎｄｅｒｓｓｏｎ，　Ｐ．Ｄｏｏｌａｎ，　Ｎ．Ｆｅｌｄｍａｎ，　Ａ．Ｆｒｅｄｅｔｔｅ，　Ｂ．Ｔｈｏｍａｓ，　ＬＤＰ　Ｓｐｅｃｉｆｉｃａｔｉｏｎ，　ｄｒａｆｔ　ｉｅｔｆ－ｍｐｌｓ－ｌｄｐ－０６．ｔｘｔ．　このＬＤＰは、以降、「ＬＤＰプロトコル」と呼ぶ。

［３］　Ｂ．Ｄａｖｉｅ，　Ｊ．Ｌａｗｒｅｎｃｅ，　Ｋ．ＭｃＣｌｏｇｈｒｉｅ，　Ｙ．Ｒｅｋｈｔｅｒ，　Ｅ．Ｒｏｓｅｎ，　Ｇ．Ｓｗａｌｌｏｗ，　Ｐ．Ｄｏｏｌａｎ，　ＭＰＬＳ　Ｕｓｉｎｇ　ＬＤＰ　ａｎｄ　ＡＴＭ　ＶＣ　Ｓｗｉｔｃｈｉｎｇ，ｄｒａｆｔ　ｉｅｔｆ－ｍｐｌｓ－ａｔｍ－０２．ｔｘｔ．

［４］　Ｂ．Ｊａｍｏｕｓｓｉ，　Ｃｏｎｓｔｒａｉｎｔ－Ｂａｓｅｄ　ＬＳＰＳｅｔｕｐ　ｕｓｉｎｇ　ＬＤＰ，　ｄｒａｆｔ－ｉｅｔｆ－ｍｐｌｓ－ｃｒ－ｌｄｐ－０１．ｔｘｔ．　このＬＤＰは、以降、「ＣＲＬＤＰ」と呼ぶ。

［５］　Ｅ．Ｂｒａｄｅｎ他，　Ｒｅｓｏｕｒｃｅ　ＲｅｓｅｒｖａｔｉｏｎＰｒｏｔｏｃｏｌ，　ＲＦＣ２２０５．　このＬＤＰは、以降、「ＲＳＶＰ」と呼ぶ。

【００２５】ノード１０は、ＳＩ連係を介してＭＰＬＳ機能性を実装する。これは、管理エンティティ、つまり管理レコードのコンテキストでＳＩを示す図４を参照することにより、よりよく理解される。ＳＩは、それと関連する内部ＩＤ番号を有する。ＡＴＭリンクレイヤ端点を表すことに加え、ＳＩは、また、レイヤ３機能性に関するＩＰアドレスも表し、ＩＰ目的にどの型のカプセル化が使用されるかも示す。また、各ＳＩは、いくつかの他の属性および方法と関連付けられる。特に、ＳＩは以下の方法または適用と関連付けることができる。（１）ＩＰ転送、（２）ＭＰＬＳ転送、（３）ＩＰ経路指定、および（４）ＭＰＬＳシグナリング。言い換えれば、ノード１０は、（１）前述のＩＰ転送手続きを介して次のホップのルータにＩＰデータパケットを転送し、（２）以下に説明するとおり、ＭＰＬＳ転送手続きを介してＩＰデータ・パケットを転送し、（３）ＩＰ経路指定プロトコルに関するメッセージを搬送するパケットを処理し、（４）ＭＰＬＳシグナリングプロトコルに関するメッセージを搬送するパケットを処理するように構成することができる。

【００２６】４．　ＭＰＬＳアーキテクチャの概要

図５は、制御カード２４のハードウェアとソフトウェアのアーキテクチャをさらに詳細に示している。ハードウェアの点では、カード２４は、複数の独立した物理的プロセッサ（図では長方形のボックスで表す）が関与する分散計算アーキテクチャを使用する。

【００２７】プロセッサ５０は、シグナリングメッセージに関するレイヤ２（「Ｌ２」）ＡＴＭアダプテーションレイヤパケットのセグメント化および再組立て機能を扱う。前述のとおり、いくつかのＳＩは、様々な型の経路指定プロトコルと関連付けられることになり、これらのＳＩのうちの１つと関連するパケットを受信した際、イングレスＩＰ転送器２２Ａは、パケット（これは、スイッチング装置２０を通過するために再セグメント化さ

れる）をＬ２プロセッサ５０に送信する。再組み立てした後、Ｌ２プロセッサ５０は、ＩＰ経路指定プロトコルと関連するシグナリングメッセージを経路指定プロセッサ５８上で実行される「ＩＰ経路指定」６８と呼ぶソフトウェアタスクに送信する。（Ｌ２プロセッサ５０とＩＰ経路指定６８の間の接続は示していない。）ＭＰＬＳ

ＬＤＰプロトコルと関連するシグナリングメッセージは、レイヤ３（Ｌ３）プロセッサ５４上で実行されるラベル管理システムタスク（ＬＭＳ）６４に送信される。ＬＭＳ６４およびＩＰ経路指定６８からの発信メッセージは、適切なイグレス回線カードおよびイグレスＳＩへのその後の送達のためにＬ２プロセッサ５０に送信される。

【００２８】ＩＰ経路指定６８は、ＩＰ内部ゲートウェイ、またはＩ－ＢＧＰ、ＩＳＩＳ、ＰＩＭ、ＲＩＰ、またはＯＳＰＦなどの経路指定プロトコルを実行する。（これらのプロトコルに関するさらなる情報に関しては、読者は、ｈｔｔｐ：ｗｗｗ．ｉｅｔｆ．ｏｒｇ．／ｈｔｍｌ．ｃｈａｒｔｅｒｓ／ｗｇ－ｄｉｒ．ｈｔｍｌを参照されたい。）これらの活動の結果、ＩＰ経路指定６８は、図６に概略で示すマスタＩＰ経路指定テーブル７５を維持する。マスタテーブル７５の各レコードは、ＩＰアドレスのためのフィールド７５ａ、宛先ＩＰアドレスまたはそのプレフィックスに対応する次のホップのルータＩＤ（これは、それ自体、ＩＰアドレスである）のためのフィールド７５ｂ、および次のホップのルータＩＤと関連するイグレスインターフェースインデックスのリスト７５ｃを含む。ネットワークトポロジが変更されると、ＩＰ経路指定は、ＩＰ転送器２２の転送テーブル３０を適切なイグレスインターフェースインデックスをそれに送信することによって更新することになる。（テーブル３０だけが、各宛先ＩＰアドレスエントリと関連する１つのイグレスインターフェースを有することを留意されたい。）

【００２９】図５に示すとおり、ノード１０は、そのそれぞれがＬＭＳ６４を含む複数のＬ３プロセッサ５４を使用する。各ＬＭＳ６４は、ＬＤＰシグナリングリンクに関するＴＣＰセッションおよびＵＤＰセッション（ＬＤＰセッション）を終了し、各ＬＳＰに関する状態マシンを実行する。下記にさらに詳細に説明するとおり、ＬＭＳ６４は、ＬＤＰセッションのセットアップおよび取外しを行う要求、およびＳＬＳＰのセットアップおよび取外しを行う要求を受信する。

【００３０】ＬＭＳ６４は、マサチューセッツ州、ＤｅｄｈａｍのＨａｒｒｉｓ　＆　Ｊｅｆｆｒｉｅｓから市販されている。相互互換性のため、ノード１０は、「翻訳」ソフトウェアであるＭＰＬＳコンテキストアプリケーションマネージャ（ＭＰＬＳ　ＣＡＭ）６５を含み、これは、ＬＭＳと制御カード２４のそれ以外のソフトウェアエンティティの間で着信または発信要求／応答を翻訳し、転送する。

【００３１】また、各Ｌ３プロセッサ５４は、呼処理タスク７２も含む。このタスクは、要求された接続に関する状態情報を維持する。

【００３２】別のプロセッサ５６が、中央ネットワーク管理システム（ＮＭＳ）（Ｎｅｗｂｒｉｄｇｅ　Ｎｅｔｗｏｒｋｓ　Ｃｏｒｐｏｒａｔｉｏｎ　４６０２０（商標）製品などの）を介して、またはネットワーク端末インターフェース（ＮＴＩ）を介してノードに直接に提供されるコマンド指示によって提示される管理要求を解釈し、それに応答することを含むユーザインターフェース機能性を提供する。ＭＰＬＳ機能性に関しては、ＰＬＳＰ、ＳＬＳＰ、およびＬＤＰセッションをプログラムする管理要求を受け入れ、それに応答するためにユーザインターフェース６６が提供される。

【００３３】ノード内で接続が確立されると、リソースの割振りおよび割振り解除を行うためにリソース制御プロセッサ５２が提供される。ＭＰＬＳ機能性に関しては、プロセッサ５２は、ＬＳＰに対して固有ラベル値を割り当てるラベル管理タスク６２を含む。

【００３４】経路指定プロセッサ５８上で、「ＭＰＬＳ経路指定」７０と呼ぶソフトウェアタスクが、ＵＩ６６と、ＩＰ経路指定６８と、Ｌ３プロセッサ５４上で実行されるＬＭＳ６４との間でインターフェースをとる。一般的に言って、ＭＰＬＳ経路指定７０はＳＬＳＰを管理する。例えば、パスセットアップ中、ＭＰＬＳ経路指定７０は、ユーザインターフェース６６からＳＬＳＰセットアップ要求を受信し、ＩＰ経路指定６８から次のホップの経路指定情報を検索し、次のホップに対するＬＤＰセッションを選択し、選択したＬＤＰセッションを使用してＳＬＳＰパスをセットアップするためにＬＭＳ６４の適切な実例を呼び出す。パスに関するラベルマッピングを受信したとき、ＬＭＳ６４は、ＭＰＬＳ経路指定７０に通知を行う。次に、ＭＰＬＳ経路指定７０が、その新しいパスのために、ＩＰ転送器２２の転送テーブル３０に対する更新を起動する。同様に、ネットワークトポロジが変更されたとき、ＭＰＬＳ経路指定７０は、これらの変更をＭＰＬＳ経路指定ドメイン内に反映させる。ＭＰＬＳ経路指定の機能が、本説明における以下の部分の焦点である。

【００３５】５．　基準ネットワーク

図７は、ルータ／ノードＡ、Ｂ、Ｃの中にＭＰＬＳ経路指定ドメインが存在し、ネットワーク８０の残りの部分が、ＯＳＰＦなどのＩＰ特定経路指定プロトコルを使用する基準ＩＰネットワーク８０を示している。ネットワークオペレータが、ノードＡから開始して、ネットワーク内のどこかに位置する宛先ＩＰアドレス１．２．３．４（以降、「ＦＥＣ　Ｚ」）に対するＳＬＳＰを確立することを所望していると想定する。（ＦＥＣは、ドラフトのＭＰＬＳ標準によれば、宛先ＩＰアドレスおよびそ

のプレフィックスを含むことを留意されたい。）ネットワークオペレータは、ノードＡで、そのＮＭＴＩまたはＮＭＳ（図示せず）を介して管理コマンドを入力し、ＦＥＣ　Ｚに対するＳＬＳＰの確立を要求することができる。使用するラベル配布プロトコルの型に応じて（例えば、ＬＤＰプロトコル、ＣＲＬＤＰ、またはＲＳＶＰ）、ネットワークオペレータは、ＳＬＳＰに対する宛先ノードを特定する、または何らかの宛先ノードまでのＳＬＳＰに関する所望の経路を明示的に特定する（すなわち、送信元経路指定ＳＬＳＰ）ことさえできる。さらなる代替手法では、ラベル配布プロトコルは、ＦＥＣＺの宛先アドレスにできる限り近接したノード（ＭＰＬＳ経路指定ドメイン内の）を識別するのに、ベストエフォートポリシー（例えば、ＬＤＰプロトコルにおける）を使用することができる。説明する基準ネットワーク内では、ノードＣが、ＦＥＣＺに対するＭＰＬＳ経路指定ドメイン内の「最も近接した」ノードであると想定する。

【００３６】ネットワーク８０内では、ノード間でＩＰ経路指定メッセージを通信するために、シグナリングリンク８２（これは、特定のＳＩと関連している）が提供される。さらに、その間でＭＰＬＳラベル配布プロトコルメッセージを通信するために、シグナリングリンク８４が提供される。各シグナリングリンク８４は、そのリンクと関連するＬＤＰセッションを有する。

【００３７】用語法として、文脈によってそうでないことが明らかな場合以外は、「イングレスＳＬＳＰ」という用語は、発信ノード（例えば、ノードＡ）でのＳＬＳＰを識別するのに使用し、「通過ＳＬＳＰ」という用語は、通過ノード（例えば、ノードＢ）でのＳＬＳＰを識別するのに使用し、また「イグレスＳＬＳＰ」という用語は、宛先ノード（例えば、ノードＣ）でのＳＬＳＰを識別するのに使用する。

【００３８】図７に示す基準ＩＰネットワークは、本発明を脈絡付けるのを助け、本発明を説明するのを助ける通常の適用例を読者に提供するために使用する。したがって、本発明は、本明細書に記載する特定の適用例によっては制限されない。

【００３９】６．　ＳＬＳＰのデータベース管理

イングレスＳＬＳＰ、通過ＳＬＳＰ、およびイグレスＳＬＳＰを作成し、監視し、また追跡するために、ＭＰＬＳ経路指定７０は、図８Ａのデータベース概略図に示すいくつかのテーブルまたはいくつかのデータリポジトリを維持する。各ＳＬＳＰはＬＤＰセッションによって管理されているので、各ノード上のＭＰＬＳ経路指定７０は、ＬＤＰシグナリングデータベース（ＬＳＬＴ）１００を使用して、そのノードとそのＬＤＰピア経路指定エンティティの間でセットアップされている様々なＬＤＰセッションを追跡する。ＬＳＬＴ１００は、ＬＤＰ経路指定ピアごとに１つのエントリまたはレコード１０４を有するハッシュテーブル１０２を含む。レコード１０４は、ハッシュテーブル１０２に対するインデックスとして機能するルータｉｄフィールド１０４ａ、およびＬＤＰセッションリスト１０６をポイントするポインタ１０４ｂ（すなわち、＊ｌｄｐ＿ｓｅｓｓｉｏｎ＿ｌｉｓｔ）を含む。ルータｉｄフィールド１０４ａは、１つまたは複数のＬＤＰセッションがそれに対して構成されているＬＤＰピアルータのＩＰアドレスを記憶する。各ＬＤＰセッションは、ポイントされたＬＤＰセッションリスト１０６内のエントリまたはレコード１０８によって代表される。ノードと所与のＭＰＬＳピアルータの間で複数のＬＤＰセッションを構成することができ、したがって、セッションリスト１０６は、複数のエントリまたはレコード１０８を有する可能性があることを留意されたい。図８Ａでは、図示するルータｉｄフィールド１０４ａによって識別されたＬＤＰピアルータに関して２つのＬＤＰセッションが構成されており、したがって、それに対応するＬＤＰセッションリスト１０６内に２つのレコード１０８が存在する。

【００４０】ＬＤＰセッションリスト１０６の各レコード１０８は、下記のフィールドを含む。

【００４１】ｉｆＩｎｄｅｘ（１０８ａ）：特定のインターフェースインデックス、およびＬＤＰアプリケーションに関して構成されているＳＩを識別するノード１０内の固有番号。図８Ｂは、このｉｆＩｎｄｅｘフィールドの構造をさらに詳細に示している。この構造は、ＳＩを担う回線カード／ＩＰモジュールに対するノード内部デバイスアドレス、イグレスポート、ＳＩ　ＩＤ番号（これは、回線カードごとにだけ固有である）、およびＬＤＰシグナリングリンクを扱う制御カード２４上のＬ３プロセッサ５４に対する識別コードまたは内部デバイスアドレスを含む。

【００４２】＊ｆｉｔ＿ｌｉｓｔ＿ｅｎｔｒｙ（１０８ｂ）：ＦＥＣ情報テーブル（ＦＩＴ）１１０に対するポインタ。下記にさらに詳細に説明するとおり、ＦＩＴは、ノードから生じるすべてのイングレスＳＬＳＰを追跡する。ｆｉｔ＿ｌｉｓｔ＿ｅｎｔｒｙポインタ１０８ｂは、このＬＤＰセッションと関連するイングレスＳＬＳＰのＦＩＴ１１０内にあるリストをポイントする。

【００４３】ｌｄｐ＿ｓｔａｔｕｓ（１０８ｃ）：ステータス指標。ステータスは、ＬＤＰセッションが使用中であるか否かを示す１ビットフラグ（図示せず）、およびそのＬＤＰセッションのためにリソースが利用可能であるかどうかを示す１ビットフラグ（図示せず）を含む。割振りのために利用可能なラベルがないとき、または関連するＳＩが動作しない状態になったとき、ＬＤＰセッションには、利用可能なリソースがないとみなされる。

【００４４】＊ｎｅｘｔ＿ｌｄｐ＿ｓｅｓｓｉｏｎ：同じＬＤＰピアルータと関連する別のＬＤＰセッションレ

コード１０８に対するポインタ
ＦＩＴ１１０は、イングレスＳＬＳＰ、すなわち、そのノードから開始したＳＬＰを追跡する。（ＦＩＴ１１０は、通過ＳＬＳＰまたはイグレスＳＬＳＰの追跡を行わないことを留意されたい。）ＳＬＳＰが構成されたとき、ＦＩＴエントリまたはレコード１１２はＭＰＬＳ経路指定７０によって作成され、ＳＬＳＰが削除されたとき、レコード１１２はＦＩＴ１００から除去される。

【００４５】各ＦＩＴエントリまたはレコード１１２は、以下の要素を含む。

【００４６】** ｐｒｅｖ＿ｆｉｔＥｎｔｒｙ（１１２ａ）：現行のエントリを参照するポインタに対するポインタ。これは、リストへの追加および除去を容易にするために使用する。

【００４７】ＦＥＣ：ＬＳＰに対するＩＰ宛先。ＦＥＣは、標準案に準拠して、宛先ＩＰアドレス１１２ｂおよびＬＳＰが宛先とするプレフィックス１１２ｃから構成される。

【００４８】Ｓｒｔ＿ｉｎｄｅｘ（１１２ｄ）：送信元－経路テーブルまたは送信元－経路リスト（ＳＲＴ）１１４に対するインデックス。このインデックスは、ＬＳＰが送信元経路指定されない場合、値０をとり、送信元経路指定される場合、０より大きい値をとる。ＳＬＳＰ確立コマンドが送信元経路指定されたパスを含む場合、ルータＩＤ　ＩＰアドレスが、図示するとおり、順番にＳＲＴ１１４内に記憶される。

【００４９】ｉｆＩｎｄｅｘ（１１２ｅ）：ＦＥＣが次のホップのルータに到達するために使用されるイグレス回線カードおよびイグレスＳＩを特定する。このフィールドの構造は、図８Ｂに示したのと同じである。ただし、ＦＩＴ１１０内で、このフィールド１１２ｅは、ＦＥＣに対するイグレスデータパス（シグナリングチャネルと対立するものとして）のためのＳＩを特定することを留意されたい。

【００５０】ｆｅｃＳｔａｔｕｓ（１１２ｆ）：ｔｔｌ値、ｉｎｇｒｅｓｓＳｅｔｕｐフラグ、ｒｅｔｒｙＳｅｑカウンタ、およびｒｅｔｒｙＳｅｃカウンタによって表される、このＦＩＴエントリの状態（図８Ｃを参照）。ｔｔｌ値は、着信パケットから減分されるべき活動時間値を示す。ｉｎｇｒｅｓｓＳｅｔｕｐフラグは、ＳＬＳＰがうまく確立されたことを示す。ｒｅｔｒｙＳｅｑカウンタは、以下にさらに詳細に説明するとおり、このＳＬＳＰをＭＰＬＳ経路指定がセットアップしようと試みた回数の記録をとる。ｒｅｔｒｙＳｅｃカウンタは、次の再試行が試みられるまでに何秒残っているかを把握する。

【００５１】ｌｓｐ＿ｉｄ（１１２ｇ）：ＭＰＬＳドメイン内のＳＬＳＰを識別するのに使用する固有識別子。本実施形態では、識別子は、ノード内のＬＳＰを一意的に識別するために、ノードＩＰルータＩＤの連結に、Ｕ１６６によって選択された固有番号を加えたものを含む。また、ｌｓｐ＿ｉｄは、ＦＩＴ１１０のためのハッシュキーとして使用する。

【００５２】＊ＲＷＰｐｔｒ（１１２ｈ）：下記にさらに詳細に説明する経路監視データベース（ＲＷＴ）１２０に対するポインタ。

【００５３】Ｎｅｘｔ．ＲＴＬＰｔｒ（１１２ｉ）、ｐｒｅｖ．ＲＴＬＰｔｒ（１１２ｊ）：ｆｅｃＳｔａｔｕｓフィールド１１２ｆのｉｎｇｒｅｓｓＳｅｔｕｐフラグが、対応するＳＬＳＰがうまくセットアップされなかったことを示すＦＩＴエントリ１１２を追跡するのに使用する順方向ポインタおよび逆方向ポインタ。これらのポインタは、ＦＩＴ１１０内に組み込まれるｒｅｔｒｙ　ｌｉｓｔ（ＲＬＴ）１１６を実装するのに、基本的に使用する。例えば、「Ａ」および「Ｂ」というラベルの付いたＦＩＴエントリ１１２が、ＲＴＬ１１６の一部を形成する。したがって、ＲＴＬは、ノードがＦＩＴ１１０を迅速に検査して、すべてのピアルータに関して待ち状態のＳＬＳＰを探し出すことができるようにする。

【００５４】＊ｎｅｘｔ＿ｆｉｔＥｎｔｒｙ（１１２ｋ）：現行のＦＥＣ／ＦＩＴエントリと同じＬＤＰセッションを使用してセットアップされた次のＦＥＣ／ＦＩＴエントリに対するポインタ。

【００５５】ＲＷＴ１２０は、ノードによって扱われるすべてのＳＬＳＰ、すなわち、イングレス、通過およびイグレスＳＬＳＰを追跡する。ＲＷＴ１２０は、宛先ＩＰアドレスフィールド１２２ａ、ＩＰプレフィックスフィールド１２２ｂ、および下記にさらに詳細に説明するＬＳＰのリスト１２４をポイントする＊ｒｗｔ－ｅｎｔｒｙ１２２Ｃを含む。

【００５６】宛先ＩＰアドレスフィールド１２２ａおよびプレフィックスフィールド１２２ｂは、使用する特定のラベル配布プロトコルに応じて、異なる型の管理エンティティを記憶するのに使用する。これらのエンティティは、（ａ）ＬＤＰプロトコルの場合、ＦＥＣ、（ｂ）非送信元経路指定ＲＳＶＰの場合、宛先ノードのルータＩＤ、（ｃ）厳密な送信元経路指定が行われるＣＲ－ＬＤＰおよびＲＳＶＰの場合、次のノードのルータＩＤ、（ｄ）ゆるい送信元経路指定が行われるＣＲ－ＬＤＰおよびＲＳＶＰの場合、構成された送信元－経路内の次のホップであることが可能である。これらはすべて、ネットワークを通ってＳＬＳＰが辿る次のホップとして要約することができる。

【００５７】テーブル１２２は、ＩＰプレフィックスフィールド１２２ｂに基づいてハッシュされることを留意されたい。通過ノードまたはイグレスノードですべてが同じＩＰプレフィックスを参照する要求されたいくつかのＳＬＳＰが存在する可能性がある。各個別ＳＬＳＰは、ＬＳＰリスト１２４内の別々のエントリまたはレコード１２６によって識別される。ただし、ノード１０上

の任意の所与のＩＰプレフィックスと関連するイングレスＳＬＳＰは、１つだけ存在することが可能である。（言い換えれば、ＬＭＳ６４から受信された次のホップの要求ごとに１つのエントリ１２６が存在し、またノード上に作成された１つのイングレスＳＬＳＰに対して１つのエントリが存在する。イグレスＳＬＳＰもまた、次のホップの情報を要求し、したがって、このテーブル内に含まれることも留意されたい。）

【００５８】各ＬＳＰリストエントリ１２６は、以下の要素を含む。

【００５９】ｐｒｅｖ＿ＲＷＴＰｔｒ（１２６ａ）、ｎｅｘｔ＿ＲｗｔＰｔｒ（１２６ｆ）：特定のＩＰプレフィックスに関する追加のエントリ１２６を追跡するのに使用する順方向ポインタおよび逆方向ポインタ。同じＩＰプレフィックス１２２ｂと関連するＬＳＰのすべては、ポインタ１２６ａおよび１２６ｆを使用してまとめてリンクする。

【００６０】ｎｅｘｔ＿ＥｇｒｅｓｓＰｔｒ（１２６ｂ）、ｐｒｅｖ＿ＥｇｒｅｓｓＰｔｒ（１２６ｃ）：下記にさらに詳細に説明するとおり、新しいＬＤＰセッションが構成されたとき、場合によっては拡張され得るイグレスＳＬＳＰを追跡するのに使用する順方向ポインタおよび逆方向ポインタ。これらのポインタは、基本的に、ＲＷＴ１２０内に組み込まれるＬＳＰイグレステーブルまたはＬＳＰイグレスリスト（ＬＥＴ）１３０を実装するのに使用する。例えば、図８Ａでは、「Ｘ」および「Ｙ」というラベルの付いたＲＷＴエントリ１２６が、ＬＥＴ１３０に属する。ＳＬＳＰをセットアップする際、対応するＦＥＣの宛先アドレスに「より近接した」さらなるＬＤＰシグナリングリンクをもはやノード１０が見つけられないときにベストエフォートポリシー（すなわち、ＬＤＰプロトコル）が使用されるときはいつでも、エントリ１２６がＬＥＴ１３０に追加される。例えば、基準ネットワーク内でＦＥＣ　Ｚに対するＳＬＳＰを確立する際、ノードＣ（ＭＰＬＳ経路指定ドメインの境界にある）が、ＦＥＣ　Ｚの宛先アドレスに向うＬＤＰシグナリングリンクを最早見つけられないとき、したがって、ノードＣが、このＳＬＳＰに関するＲＷＴエントリ１２６を作成するとき、エントリがＬＥＴに追加される。

【００６１】ｆｉｔＥｎｔｒｙＰｔｒ（１２６ｄ）：このＲＷＴエントリ１２６に対応するＦＩＴエントリ１１２に対するポインタ。このフィールドの値は、このノードで作成されたイングレスＳＬＳＰに対するものを除き、すべてのエントリに関して零となる。

【００６２】Ｌ３＿ｉｄ（１２６ｅ）：ＬＳＰに関する次のホップの要求を最初に要求したＬ３プロセッサのアドレスまたは識別、またはイングレスＳＬＳＰをセットアップするのに使用するＬ３プロセッサのアドレスまたは識別。

【００６３】ｌｓｐ＿ｉｄ（１２６ｇ）：ＦＩＴ１１０内のｌｓｐ＿ｉｄ１１２ｇと同じであるが、これらのＬＳＰは、他のノードで開始された可能性があることが異なる。

【００６４】７．　ＬＤＰセッションの確立

ＬＤＰセッションは、ＵＩ６６を介して受信され、ＭＰＬＳ経路指定７０に転送される管理要求を介して構成される。ＵＩ６６によって得られるデータには、ＬＤＰシグナリングリンクＳＩのＡＴＭリンクレイヤ端点（すなわち、回線カードアドレス、ポート、ＶＰＩ／ＶＣＩ）、ＳＩに割り当てられたＩＰアドレス、ならびにラベル範囲、ラベル空間ＩＤ、およびキープアライブタイムアウトなどのＬＤＰ特定パラメータが含まれる。

【００６５】ＭＰＬＳ経路指定７０は、ラウンドロビンアルゴリズムを使用して、ＬＭＳ６４の１つの実例（すなわち、Ｌ３プロセッサ５４のうちの１つ）を選択し、新しいＬＤＰセッションを確立するように関連するＭＰＬＳ　ＣＡＭ６５に要求する。ＭＰＬＳ　ＣＡＭは、ネットワークオペレータによって選択されたＳＩ上でＬＤＰシグナリングアプリケーションを使用可能にし、Ｌ２プロセッサ５０と関連するフィルタリング機構（図示せず）を含むノードを構成し、回線カード１２と選択されたＬＭＳ／Ｌ３プロセッサ５４の間で、特定のＬＤＰシグナリングＳＩと関連するすべてのＬＤＰパケットが伝搬（イングレス方向とイグレス方向の両方で）され得るようにする。これが行われると、ＬＭＳ６４は、該当するラベル配布プロトコル（例えば、ＬＤＰプロトコル、ＣＲＬＤＰ、ＲＳＶＰ）に従って、ＬＤＰピアルータに対してＬＤＰセッション確立メッセージを送り出す。これらのメッセージには、「ハロー」およびその他のセッション確立メッセージが含まれる。

【００６６】ＬＤＰピアルータとのＬＤＰセッションが確立されると、ＬＭＳ６４は、ラベルマネージャ６２にＬＤＰセッションに関するラベル交渉範囲（標準案によればＬＤＰセッションを確立する機能の１つ）を通知する。また、ＬＭＳ６４は、ＬＤＰピアルータのＩＰアドレスをＭＰＬＳ経路指定７０に渡し、ＭＰＬＳ経路指定７０は、このアドレスをＬＳＬＴ１００のルータＩＤフィールド１０４ａ内に記憶する。さらに、ＬＭＳ６４は、ＬＤＰシグナリングＳＩを識別するインターフェースインデックスをＭＰＬＳ経路指定７０に渡し、ＭＰＬＳ経路指定７０は、それをＬＳＬＴ１００のｉｆＩｎｄｅｘフィールド１０８ａ内に記憶する。

【００６７】８．　ＳＬＳＰの確立

８．１　イングレスノードでの手続き

基準ネットワークを参照すると、ＦＥＣ　Ｚに対してノードＡにおいてＳＬＳＰが、ノードＡのＮＭＴＩまたはノードＡと通信するＮＭＳを介してネットワークオペレータによって明示的に確立されなければならない。ＳＬＳＰを構成する命令は、パラメータの１つとしてＺ、す

なわち、ＦＥＣ　Ｚに関する宛先ＩＰアドレスおよびそのプレフィックスを含む。コマンドは、ＵＩ６６によって受信されて解釈される。

【００６８】ＵＩ６６は、固有のＬＳＰ　ＩＤを選択し、これは、前述のとおり、好ましくは、ノードのＩＰルータＩＤおよび固有番号の連結を含む。次に、ＵＩ６６は、ＦＥＣ　Ｚに対してＳＬＳＰを作成し、それを選択したＬＳＰ　ＩＤと関連付けるようにＭＰＬＳ経路指定７０に要求する。

【００６９】ＭＰＬＳ経路指定７０は、ＩＰ経路指定６８からのＦＥＣ　Ｚに関する次のホップ情報を要求する。これは、次のホップの情報を得るために非送信元経路指定ＬＳＰに関して行われ、また送信元経路指定内の情報（ネットワークオペレータによって供給される）を検証するために、送信元経路指定ＬＳＰに関しても行われる。より具体的には、ＭＰＬＳ経路指定７０は、次の手続きを実行してこの新しいＦＥＣに関するＳＬＳＰの確立を開始する。

【００７０】図９をさらに参照すると、第１ステップ１５０で、ＭＰＬＳ経路指定７０が、ＦＥＣ　Ｚと同じ宛先ＩＰアドレスおよび同じプレフィックスを有する既存のエントリ１１２を求めてＦＩＴ１１０を探索する。次にステップ１５２で、ＦＩＴ１１０内にそのようなエントリが存在する場合、ＭＰＬＳ経路指定７０は、このノードからＦＥＣ　Ｚが既に確立されていることを示す障害コードを戻す。ステップ１５８で、ＭＰＬＳ経路指定７０は、新しいＦＩＴエントリ１１２を作成し、それをＦＩＴ１１０に追加する。また、対応するエントリ１２６も、ＲＷＴハッシュテーブル１２２内のＦＥＣ　Ｚに関するＬＳＰリスト１２４内に挿入される。必要があれば、ＭＰＬＳ経路指定７０は、ＦＥＣ　ＺのＩＰプレフィックスおよびＩＰアドレス、または明示的経路内の第１のホップのＩＰプレフィックスおよびＩＰアドレスを含む新しいエントリ１２２をＲＷＴ１２０に追加する。

【００７１】ステップ１６０で、ＭＰＬＳ経路指定７０は、ＦＥＣ　Ｚに到達するための次のホップに関するピアＩＰアドレス（または非送信元経路指定ＲＳＶＰの場合、宛先ノードのルータｉｄ、またはゆるい送信元経路指定が行われるＣＲ－ＬＤＰおよびＲＳＶＰの場合、構成された送信元経路指定内の次のホップ）を提供するようにＩＰ経路指定６８に要求する。これが得られると、ステップ１６２で、ＭＰＬＳ経路指定７０は、次のホップのルータＩＤに一致するＬＳＬＴエントリ１０２を探索する。一致するＬＳＬＴエントリが存在する場合には、ステップ１６４で、ＭＰＬＳ経路指定７０は、対応するＬＤＰセッションリスト１０６から利用可能なＬＤＰセッションを選択する。これは、循環リンクされたリストであり、ＬＳＬＴエントリ１０２内の＊ｌｄｐ＿ｓｅｓｓｉｏｎ＿ｌｉｓｔポインタ１０４ｂが、ＭＰＬＳ経路指定７０によって選択される次のＳＬＳＰセットアップのために使用されるＬＤＰセッションをポイントするように管理される。ＬＤＰセッションが選択されると、ＦＥＣ　Ｚに関して新しく作成されたＦＩＴエントリ１１２が、同じＬＤＰセッションを使用してその他のＦＩＴエントリにリンク（＊＊ｐｒｅｖ＿ｆｉｔＥｎｔｒｙポインタ１１２ａおよび＊ｎｅｘｔ－ＦｉｔＥｎｔｒｙポインタ１１２ｉを介して）される。

【００７２】＊ｎｅｘｔ＿ｌｄｐ＿ｓｅｓｓｉｏｎポインタ１０８ｄは、ＬＤＰセッションリスト内の次のセッションをポイントする。（リスト内に１つのＬＤＰセッションだけが存在する場合には、＊ｎｅｘｔ＿ｌｄｐ＿ｓｅｓｓｉｏｎは自らをポイントする。）ＦＩＴ１１０とＬＤＰセッションリスト１０６の間のリンクが作成されると、ＭＰＬＳ経路指定７０は、リソースを有するＬＤＰセッションリスト内の次のセッションをポイントするように＊ｌｄｐ＿ｓｅｓｓｉｏｎ＿ｌｉｓｔポインタ１０４ｂを更新する。これは、所与のＦＥＣに対するＬＤＰセッションを選択することに対するラウンドロビン手法をもたらす。リソースを有するピアＬＤＰルータに対するセッションが存在しない場合、＊ｌｄｐ＿ｓｅｓｓｉｏｎ＿ｌｉｓｔポインタ１０４ｂは更新されない。この場合、リストは、パスがセットアップされた後、ＭＰＬＳ経路指定７０がセッションを探すのをやめる前に、一度、検査される。

【００７３】また、ＭＰＬＳ経路指定７０が、次のホップのルータＩＤに一致するＬＳＬＴエントリ１０２を見つけられなかった場合には、対応するＬＤＰシグナリングリンクは存在しないことも留意されたい。この場合、ＭＰＬＳ経路指定７０は、ＦＥＣ　Ｚに関して新しく作成されたＦＩＴエントリをステップ１６６でＲＴＬに追加して、ステップ１６８で、適切な障害コードを持って戻る。

【００７４】ＳＬＳＰの確立を通知するＬＤＰセッションが選択されると、次にステップ１７０で、ＭＰＬＳ経路指定７０は、ＳＬＳＰのセットアップを通知するようにＬＭＳ６４に要求する。ノードＡのＬＭＳ６４は、ＬＤＰ標準案に従い、ラベル要求メッセージをその下流のＬＤＰピアルータであるノードＢに送信し、ＦＥＣＺに対するＬＳＰのセットアップを所望することを示す。ラベル要求メッセージは、経路指定プロトコルに従い（ホップごとに、または送信元経路指定されて）ＭＰＬＳ経路指定ドメインを介して下流に向ってイグレスノードＣまで伝播され、またラベルマッピングメッセージは、逆にイグレスノードＡに戻るように上流に伝播される。最終的に、図１０に示すとおり、ラベルメッセージは、ＦＥＣＺに対してＭＰＬＳ経路指定７０によって選択されたＬＤＰシグナリングリンク上で帯域内で受信されなければならない。このラベルメッセージは、ラベル、すなわち、ＩＰパケットおよびそのＡＴＭセルをノードＢに転送するのに使用すべきＶＰＩ／ＶＣＩ値を識別する。

ラベルは、ＭＰＬＳ経路指定７０およびラベルマネージャ６２に渡される。さらに、ステップ１７４で、ＬＭＳ６４が、データトラフィックを扱うためにイグレス回線カードおよびポート上で使用されるＳＩに対するイグレスインターフェースインデックスを構成するように、呼プロセッサ７２に通知する。（イグレス回線カードは、ＦＥＣ　Ｚに関するＳＩを通知するＬＤＰと関連する回線カードおよびポートと同じものであることを留意されたい。）これは、ＦＥＣ　ＺをＡＴＭ　ＶＰＩ／ＶＣＩラベルに「バインド」する。この結付けは、ＭＰＬＳ経路指定７０に報告され、ＭＰＬＳ経路指定７０は、ステップ１７６で、ＦＥＣ　Ｚに一致するエントリ１１２を求めてＦＩＴ１１０を探索し、その時点で、呼プロセッサ７２から得られたイグレスインターフェースインデックスを使用してｉｆＩｎｄｅｘフィールド１１２ｅが更新される。

【００７５】さらに、ＭＰＬＳ経路指定７０は、ｒｅｔｒｙＳｅｑカウンタおよびｒｅｔｒｙＳｅｃカウンタをゼロに設定することによってｆｅｃＳｔａｔｕｓフィールド１１２ｆ（図８Ｃ）を更新し、またｉｎｇｒｅｓｓＳｅｔｕｐフラグを１に設定して、セットアップが成功したことを示す。ステップ１７８で、ＭＰＬＳ経路指定７０は、新しく確立されたＳＬＳＰおよびそのイグレスインターフェースインデックスに関してＩＰ経路指定６８に通知を行い、その時点で後者のタスクは、そのＩＰ転送テーブル７５（図６）を更新して、新しく確立されたイグレスインターフェースインデックス（参照番号７６で概略を示す）を適切なリスト７５ｃに追加する。ＩＰ経路指定６８の方は、リスト７５ｃ内にいくつかの潜在的なイグレスインターフェースインデックスを有する可能性があり、パケットを転送するのに使用することができる。これらの選択肢の中でどれかに決めるため、ＩＰ経路指定６８は、ＭＰＬＳによって使用可能になったイグレスインターフェースインデックス（ＦＥＣごとに１つだけ存在することが可能である）に、非ＭＰＬＳイグレスインターフェースよりも高い優先順位を与える優先順位スキームを使用する。優先順位スキームは、ビットマップ７５ｄ（１つだけを示している）の機構を介して実行され、このビットマップは、イグレスインターフェースインデックスリスト７５ｃの各エントリと関連している。ビットマップ７５ｃは、どの型のアプリケーション、例えば、ＳＬＳＰまたはＩＰが、イグレスインターフェースインデックスエントリと関連しているかを示す。この優先順位スキームに従って、ステップ１８０で、ＩＰ経路指定は、各ＩＰ転送モジュールの転送テーブル３０に、新しく作成されたイグレスインターフェースインデックス７６をダウンロードする。（テーブル３０は、各ＩＰアドレスまたはそのプレフィックスごとに単一のイグレスインターフェースインデックスだけをリストする。）非同期で、ＭＰＬＳ経路指定７０も、ステップ１８２で、ＦＥＣ　Ｚに対するイングレスＳＬＳＰがうまく作成されたことをＵＩ６６に通知する。

【００７６】所定期間内にラベルマッピングメッセージが全く受信されなかった場合、またはノードＢから受信されたシグナリングメッセージが、ＦＥＣ　Ｚに対するＳＬＳＰのセットアップを拒否する場合、ＬＭＳ６４は、ステップ１８４で障害をＭＰＬＳ経路指定７０に通知する。ＭＰＬＳ経路指定は、これに伴い、ＦＥＣ　Ｚに関するＦＩＴエントリ１１２をＲＴＬ１１６上に配置し、ｆｅｃＳｔａｔｕｓｉｎｇｒｅｓｓＳｅｔｕｐフィールド（図８Ｃ）をゼロに設定して、ｒｅｔｒｙＳｅｑフィールドの値を増分する（最大で６まで）。ステップ１８６で、ＭＰＬＳ経路指定は、障害をＵＩ６６に通知する。

【００７７】ＦＩＴエントリに関する再試行機構は、ＳＬＳＰパスセットアップが、１０秒、２０秒、３０秒、４０秒、５０秒、６０秒で再試行されるようにする線形バックオフ機構である。毎１０秒ごとに作動するＭＰＬＳ経路指定７０に関連する１つの再試行タイマが存在する。この時点で、ＭＰＬＳ経路指定は、ＲＴＬ１１６を検査して、ＲＴＬ１１６内の各ＦＩＴエントリ１１２に残された時間量（図８Ｃのｒｅｔｒｙ Ｓｅｃ）を減分する。ｒｅｔｒｙＳｅｃ値がゼロの場合、ＦＩＴエントリ１２は、ＲＴＬ１１６から除去され、再試行順序番号が１だけ増分されて、イングレスＳＬＳＰを確立しようとする新しい試行が行われる。再試行が成功した場合、ｒｅｔｒｙＳｅｑはゼロに設定され、ｉｎｇｒｅｓｓＳｅｔｕｐフラグは１に設定される。再試行が失敗した場合には、ＦＩＴエントリは再びＲＴＬに加えられ、ｒｅｔｒｙＳｅｑが増分される（最大順序番号は、好ましくは、６である）。ｒｅｔｒｙＳｅｑが増加されたとき、ＭＰＬＳ経路指定７０がＳＬＳＰのセットアップを再試行する期間も、次に大きい間隔に増加される。例えば、ｒｅｔｒｙＳｅｑが２から３に増加したとき、再試行間の時間間隔は、２０秒から３０秒に増加する、すなわち、ｒｅｔｒｙＳｅｃは３０に設定される。ｒｅｔｒｙＳｅｑが６に等しいとき、再試行間の間隔は６０秒である。

【００７８】８．２　通過ノードでの手続き

通過ノードＢで、ＦＥＣ　Ｚに関するラベル要求メッセージが、ＭＰＬＳシグナリングリンク８４上で受信され、Ｌ２プロセッサ５０によって担当のＬＭＳ６４に転送される。ＬＭＳ６４は、次のホップの情報をＭＰＬＳ経路指定７０から要求し、ＭＰＬＳ経路指定７０の方は、ＦＥＣ　Ｚに関する次のホップのルータＩＤをＩＰ経路指定６８から検索して、次のホップのルータＩＤをＲＷＴ１２０内に記憶し、次のホップのＬＤＰピアルータであるノードＣに対する下流のＬＤＰセッションを選択し、前述のとおり、このデータをＬＭＳ６４に供給する。次に、ＬＭＳ６４は、ラベル交渉範囲（上流のノー

ドＡとのＬＤＰセッションが確立されたときに決まる）の中からＶＩＰ／ＶＣＩラベルを予約するようにラベルマネージャ６２に要求する。このラベルは、ノードＡに上流に向って、ラベルマッピングメッセージが送信されるときに転送される。次に、必要な場合、上流のラベル要求メッセージを受信したＬＭＳ６４は、そのラベル要求メッセージをノードＣまで進めるために、下流のＬＤＰセッションを担うＬＭＳ（異なるＬ３プロセッサ５４上の）の別の実例を通知する。

【００７９】ラベルマッピングメッセージを下流のシグナリングリンクから受信したとき、ＬＭＳ６４は、ラベル、すなわち、上流のノードＡと関連するＶＰＩ／ＶＣＩとラベル、すなわち、下流のノードＣと関連するＶＰＩ／ＶＣＩの間のクロスコネクトを確立するように呼プロセッサ７２に通知し、下流へのデータフローを確立する。通過ノードでは、これは、前述したとおり、ＡＴＭ型のクロスコネクトをもたらす。さらに、ノードＡに対する上流のＬＤＰセッションを担うＬＭＳ６４は、ラベルマネージャ６２によって事前に予約されたラベルとともにラベルマッピングメッセージをノードＡに転送する。

【００８０】送信元経路指定ＳＬＳＰの場合、通過ノードＢがＩＰ経路指定７０から次のホップの情報を得る必要はない可能性があることを留意されたい。ただし、これは、送信元経路指定リスト内に提供される次のホップが正確であることを通過ノードがその内部経路指定テーブルを介して確認できるようにする（例えば、次のホップが、要求された宛先ＩＰアドレスまたはプレフィックスの下にリストされているかどうかを検査することにより）好ましい機能である。明示的に経路指定された次のホップが確認できない場合には、エラーを宣言することができる。

【００８１】８．３　イグレスノード上の手続き

イグレスノードＣ上では、ラベル要求メッセージが、ノードＢとの上流のシグナリングリンク上で受信され、Ｌ２プロセッサ５０によって担当のＬＭＳ６４に転送される。ＬＭＳ６４は、次のホップの情報をＭＰＬＳ経路指定７０から要求し、ＭＰＬＳ経路指定７０の方は、ＩＰ経路指定６８から次のホップの情報を要求する。ただし、この場合、次の状況のうちの１つが生じる。（１）ＩＰ経路指定６８によって戻される次のホップのルータＩＤは、現行のノードである、または（２）次のホップは見つかるが、次のホップまでのＬＤＰセッションが存在しない（すなわち、ＭＰＬＳドメインの端に達している）。これらの場合のどちらでも、ＭＰＬＳ経路指定７０は、ＦＥＣ　Ｚに対するＳＬＳＰがこのノードで出なければならないことをＬＭＳ６４に通知し、ＬＭＳ６４は、前述のとおり、上流のノードＢにラベルマッピングメッセージを送信するが、ＦＥＣ　Ｚに関するラベル要求メッセージは先に進めない（またそうすることができない）。この場合、ＭＰＬＳ経路指定７０は、前述のとおり、エントリ１２６をＲＷＴ１２０に追加するが、新しく作成したＲＷＴエントリ１２６もＬＥＴ１３０に追加する。

【００８２】この場合、ＬＭＳ６４は、ＩＰ転送のために構成されたＳＩを確立するように呼プロセッサ７２に指示する。このＳＩは、ＳＬＳＰに関してノードＢとＣの間でＭＰＬＳラベルとして使用されるＶＰＩ／ＶＣＩと等しいＡＴＭ端点（すなわち、ＶＰＩ／ＶＣＩ）を有する。

【００８３】９．　スイッチング／経路指定活動

ＦＥＣ　Ｚに対するＳＬＳＰのセットアップは説明したので、次に、ＦＥＣＺに関連するＩＰパケットが処理される仕方を簡単に説明する。イングレスノードＡで、ＩＰパケットは、複数のＡＴＭセルの形式でポート１４Ａ１に着信し、これをＩＰ転送器２２Ａが、構成ＩＰパケットに再組立てする。受信パケットの宛先ＩＰアドレスが分かると、ＩＰ転送器２２Ａは、「最も近接した」エントリを求めて転送テーブル３０を検査する。これが、ＦＥＣ　Ｚに対するＳＬＳＰの確立に関連してＩＰ経路指定６８によってダウンロードされたＦＥＣ　Ｚに対するエントリとなる。したがって、転送テーブル３０は、イグレス回線カード１２Ｂの識別またはアドレス、イグレスポート１４Ｂ１、およびイグレスＳＩ番号を含むイグレスインターフェースインデックス７６を提供する。イグレスインターフェースインデックスは、パケットに付加される。また、イングレスＩＰ転送器２２ＡはＴＩ１８を選択し、部分的にはパケット内にカプセル化されたＣＯＳフィールドに基づき、スイッチング装置２０を介してパケットをイグレスＩＰ転送器２２Ｂに伝送する。次に、パケットは、スイッチング装置２０を介する選択されたＴＩ１８上での伝送のために再セグメント化され、イグレスＩＰ転送器２２Ｂによって受信される。イグレスＩＰ転送器２２Ｂの方は、パケットに付加されたイグレスＳＩおよびＣＯＳ情報を抽出し、イグレスインターフェースインデックス（すなわち、イグレスＳＩ）によって示されるカプセル化と一致するように変更する。これには、パケットにＶＰＩ／ＶＣＩラベルを付加することが含まれる。パケットは、その後にＡＴＭを構成するセルにセグメント化され、イグレスＳＩによって示されるＶＰＩ／ＶＣＩ値とともにイグレスポート１４ＢＩから送信される。

【００８４】通過ノードＢ上で、ＩＰパケットに対応するＡＴＭセルがイングレスポートによって受信される。ＣＡＭ１５が、ＡＴＭイグレス接続インデックスを戻し、セルがＡＴＭセルとして処理される。また、イングレス回線カード１２Ａは、ＣＡＭ１５Ａから検索した内部アドレス指定情報を各セルに付加し、セルをイグレス回線カードに経路指定できるようにし、このイグレス回線カードが、セルのＶＰＩ／ＶＣＩ値を置き換える。次

に、イグレス回線カードは、この新しいＶＰＩ／ＶＣＩ値を使用してセルを送信する。この場合、ＩＰ転送モジュール２２は、スイッチング動作に積極的には関与しておらず、ＩＰパケットを再組立てまたは再セグメント化する、またはＩＰ経路指定ルックアップを行う必要がないことを留意されたい。

【００８５】イグレスノードＣ上で、ＩＰパケットに対応するＡＴＭセルが、イングレスポートによって受信され、セルによって搬送されたＶＰＩ／ＶＣＩに合せて構成されたＳＩに従って処理される。ＳＩは、セルがＩＰ転送モジュール２２Ａに送信され、より高いレイヤのＩＰパケットに再組立てが行われ、その後、通常のＩＰパケットとして処理されるように構成されている。

【００８６】１０．　ネットワークトポロジ変更

１０．１　新しいＬＤＰセッション

新しいＬＤＰセッションがノード１０上で確立されたとき、ＬＭＳ６４は、このイベントについてＭＰＬＳ経路指定７０に通知し、新しいＬＤＰセッションに関するインターフェースインデックスについて知らせる。この信号は、ノードが新しいＬＤＰセッションの開始側であっても、応答側であっても生じる。図１１の流れ図をさらに参照すると、ステップ１９０で、ＭＰＬＳ経路指定が、ＬＳＬＴ１００内でピアルータＩＤ　ＩＰアドレスを探索する。このルータに対するＬＳＬＴエントリ１９４が見つかった場合には、ステップ１９２で、ＭＰＬＳ経路指定７０が対応するＬＤＰセッションリスト１０６を検査し、新しいＬＤＰセッションと同じインターフェースインデックスを有するＬＤＰセッションに対するエントリ１０８が存在しないことを確認する。そのようなエントリが見つからなかった場合、ステップ１９５で新しいエントリ１０８が作成される。そのようなエントリが見つかった場合、エラーが戻される。新しく構成されたＬＤＰセッションに対するピアルータＩＤに一致するＬＳＬＴエントリ１０４が見つからなかった場合には、ステップ１９４で、ＭＰＬＳ経路指定が、新しいＬＳＬＴエントリ１０４を作成して、それを挿入し、その後、ステップ１９５でＬＤＰセッションリストエントリ１０６が作成される。

【００８７】ステップ１９６で、ＭＰＬＳ経路指定７０が、ＬＥＴ１３０を検査する。ＬＥＴに属する各ＲＷＴエントリ１２６に対し、対応するＦＥＣがハッシュテーブル１２２から決定され、ステップ２００で、そのＦＥＣに対する次のホップのルータＩＤがＩＰ経路指定６８から要求される。ステップ２０１で、次のホップのルータＩＤが、新しく構成されたＬＤＰセッションのピアルータＩＤに対して比較される。一致が見つからなかった場合、制御は、ステップ１９８に戻り、また一致が見つかった場合、制御は、ステップ２０２に移行する。ステップ２０２で、ＭＰＬＳ経路指定７０は、識別されたＦＥＣのための新たに到達可能なピアルータにラベル要求メッセージを送信するようにＬＭＳ６４に指示する。

【００８８】１０．２　シグナリングリンク障害

ノード上でＬＤＰセッションに障害が起きたとき、ノードは、関連するＶＰＩ／ＶＣＩ範囲を使用するすべてのＳＬＳＰ（ラベルマネージャ６２内に記憶された）を転送するのを停止して、ノードからのクロスコネクトを除去する。また、ノードは、障害の起きたＬＤＰセッションと関連する各ＳＬＳＰの上流のピアにラベル撤回メッセージを送信する。例えば、ＭＰＬＳリンク８４ＢＣ（図７）に障害が起きた場合、ノードＢは、ＦＥＣ　Ｚに関するラベル撤回をイングレスノードＡに送信する。ラベル撤回メッセージがイングレスノードＡで受信されたとき、ノードＡは、そのパスを使用することを停止し（代りに、ＩＰホップごとの転送が使用される）、前述したステップを即時に再開してＦＥＣ　Ｚに対するパスを再確立する。これが成功しなかった場合には、ＦＥＣ　Ｚに対するＳＬＳＰがＲＴＬ１１６上に配置され、その後、前述した再試行手続きが実行される。

【００８９】さらに、ＬＤＰセッションが、何らかの理由で、イングレスノードＡ内で動作しなくなった場合、ＬＭＳ６４はＭＰＬＳ経路指定７０に知らせる。この呼の一環として、ＬＭＳ６４は、ＭＰＬＳ経路指定７０にピアルータＩＤ　ＩＰアドレスを提供する。次に、ＭＰＬＳ経路指定７０は、ＬＳＬＴ１００のルータＩＤフィールド１０４ａ内でピアＩＰアドレスを探索する。ピアＩＰアドレスに対するエントリが存在しない場合、エラーが戻される。ピアＩＰアドレスに対するエントリ１０４ａが存在する場合、対応するセッションリスト１０６で障害の起きたＬＤＰセッションが探索される。一致するＬＤＰセッションエントリ１０８が存在する場合、そのエントリがセッションリスト１０６から除去される。

【００９０】除去されたセッションリストエントリ１０６の* ｆｉｔ＿ｌｉｓｔ＿ｅｎｔｒｙポインタ１０８ｂは、障害の起きたＬＤＰセッションを使用するすべてのイングレスＳＬＳＰを代表するすべてのＦＩＴエントリ１１２のリストをポイントする。これらのエントリのそれぞれに対して、ＭＰＬＳ経路指定７０は、前述のとおり、即時にイングレスＳＬＳＰを再確立しようと試みて、イングレスＳＬＳＰをセットアップするのに使用することができる代替のＬＤＰセッションが存在するかを調べる。再試行が成功しなかった場合、イングレスＳＬＳＰは、ＲＴＬ１１６上を進み、先に概略を述べた再試行手続きが行われる。

【００９１】１０．３　ＩＰ経路指定変更

時間の経過とともに、ＩＰ経路指定６８は、ＦＥＣ　Ｚに対する新しい次のホップを発見することが可能である。例えば、基準ネットワーク内では、ノードＢ上のＩＰ経路指定は、ＦＥＣ　Ｚに対する次のホップがノードＤ（図示せず）であるべきことを発見することができる。そのような発見をすると、ノードＢ上のＩＰ経路指

定６８は、ＦＥＣ　Ｚに関する新しい次のホップのルータＩＤをＭＰＬＳ経路指定７０に通知する。ＭＰＬＳ経路指定７０は、以下のプロセスを使用して、ＦＥＣ　Ｚに対するＳＬＳＰを再経路指定する。第１に、ＭＰＬＳ経路指定は、ＩＰプレフィックスアドレスに一致するＲＷＴエントリ、例えば、ＩＰ経路指定テーブル７５内で変更されたＦＥＣ　Ｚを探索する。エントリ１２２が見つからなかった場合、ＭＰＬＳ経路指定は戻り、見つかった場合は継続して、次に、新しい次のホップＤのルータＩＤに一致するＬＳＬＴエントリ１０４を探索する。ＬＳＴＬエントリ１０４が存在する場合、したがって、新しいルータＤに対するＬＤＰセッションが存在する場合、ＭＰＬＳ経路指定は、ルータＤに対するＬＤＰセッションを使用して、一致するＲＷＴエントリ１２２によってポイントされるＲＷＴリスト１２４内の各通過ＳＬＳＰを進むようにＬＭＳ６４に要求する。したがって、通過ＳＬＳＰは、新しい次のホップのルータＤに再経路指定される。ただし、新しい次のホップのルータＩＤに対するＬＳＬＴエントリ１０２が存在しない場合、したがって、それに対するＬＤＰセッションが存在しない場合には、ＭＰＬＳ経路指定７０は、古いホップのルータに対応するＲＷＴリスト１２４内の各通過ＳＬＳＰをＬＥＴ１３０上に配置し、そのようなＳＬＳＰをイグレスＳＬＰとみなすべきことをＬＭＳ６４に通知する。ＬＭＳ６４の方は、イグレスＳＬＳＰに対するイグレスＳＩセットアップするように呼プロセッサ７２に指示する。

【００９２】また、ＭＰＬＳ経路指定７０は、影響を受けたＦＥＣに一致するＦＩＴエントリ１１２を探索する。そのＦＥＣに一致するＦＩＴエントリ１１２が存在し、ｆｅｃ＿ｓｔａｔｕｓフィールド１１２ｆのｉｎｇｒｅｓｓ＿ｓｅｔｕｐ　フラグがゼロでない（すなわち、パスがセットアップされている）場合、ＭＰＬＳ経路指定７０は、下流のルータにラベル開放メッセージを送信することにより、ＬＭＳ６４がイングレスＳＬＳＰを閉じることを要求する。次に、ＭＰＬＳ経路指定７０は、新しい次のホップに対するルータＩＤ　ＩＰアドレスに一致するＬＳＬＴエントリ１０４ａを探索する。そのようなＬＳＬＴエントリが存在する場合には、ＬＤＰセッションが対応するＬＤＰセッションリスト１０６から選択され、前述のとおり、イングレスＳＬＳＰを確立するための手続きが行われる。

【００９３】１０．４　物理リンク障害

２つのノード間の物理リンクに障害が起きた場合には、ＭＰＬＳシグナリングとＩＰ経路指定の両方に関するシグナリングリンク８２および８４（図７参照）に障害が起きる。本実施形態では、ＩＰ経路指定６８は、リンクがダウンしていることに気付き、そのリンク上のどのＬＤＰセッションもダウンしていることにＬＭＳ６４が気付く前にその経路指定テーブル７５を更新する。これは、ＬＤＰセッションおよびＩＰ経路指定内のシグナリングセッションに対する「タイムアウト」期間を適切に設定して、インターフェースの障害が、ＭＰＬＳ経路指定７０よりもずっと迅速にＩＰ経路指定６８内に反映されるようにすることによって実現される。したがって、ＩＰ経路指定６８は、影響を受けたＳＬＳＰに関する新しい次のホップのルータＩＤについてＭＰＬＳ経路指定７０に通知し、また前述したとおり、ＭＰＬＳ経路指定７０は、ＩＰ経路指定６８によって識別された新しい次のホップのルータを使用して、現行のノードからこれらのＳＬＳＰパスを再経路指定する。これは、シグナリングリンクがダウンしていることをＭＰＬＳ経路指定７０が気付いた場合に行われるように、影響を受けたＳＬＳＰを取り外してイングレスノードに戻し、それを再通知することよりも効率的である。

【００９４】以上の実施形態は、説明の目的で、ある程度、特定的なものとして記載した。本発明の趣旨および範囲を逸脱することなく、本明細書に開示する実施形態に多数の変形および変更を加えるのが可能なことが、当分野の技術者には理解されよう。

【図面の簡単な説明】

【図１】ＡＴＭセルおよびＩＰパケットを処理するネットワークノードのシステムを示すブロック図である。

【図２】図１のノード内でどのようにＩＰパケットが処理されるかを示すプロセス流れ図である。

【図３】図１のノードの入力／出力コントローラと関連するＩＰ転送器によって使用される転送テーブルを示す図である。

【図４】図１に示すようなノードと関連する「サービスインターフェース」を表すデータ構造を示す図である。

【図５】図１のノード上の制御カードと関連するハードウェアプロセッサおよびソフトウェアプロセスのアーキテクチャを示すブロック図である。

【図６】ＩＰネットワークと関連するマスタＩＰ経路指定テーブルを示す図である。

【図７】ＩＰネットワーク内のＭＰＬＳドメインを示す基準ネットワークの図である。

【図８Ａ】通知されたラベルスイッチパス（ＳＬＳＰ）を管理するために図１のノードによって使用されるデータベースの概略図である。

【図８Ｂ】図８Ａのデータベースのあるフィールドをより詳細に示す図である。

【図８Ｃ】図８Ａのデータベースのあるフィールドをより詳細に示す図である。

【図９】ＳＬＳＰを確立する際に図１のノードによって実行されるステップを示す論理流れ図である。

【図１０】ＳＬＳＰを確立する際に図１のノードによって実行されるステップを示す論理流れ図である。

【図１１】新しいＳＬＳＰシグナリングリンクが確立された場合、ノードによって実行されるステップを示す論理流れ図である。

【符号の説明】

１０　二重機能のＡＴＭスイッチおよびＩＰルータ
１２Ａ、１２Ｂ、１２Ｃ　回線カード
１４Ａ１、１４Ａ２、１４Ｂ１、１４Ｂ２、１４Ｃ１、
１４Ｃ２　ポート
１５Ａ、１５Ｂ、１５Ｃ　ＣＡＭ
１６Ａ、１６Ｂ　トランスポート段階
１８　トランスポートインターフェース
２０　スイッチング装置
２１　接続メッシュ
２２Ａ、２２Ｂ、２２Ｃ　ＩＰ転送器
２４　制御カード
３０　ＩＰ転送テーブル

【図１】

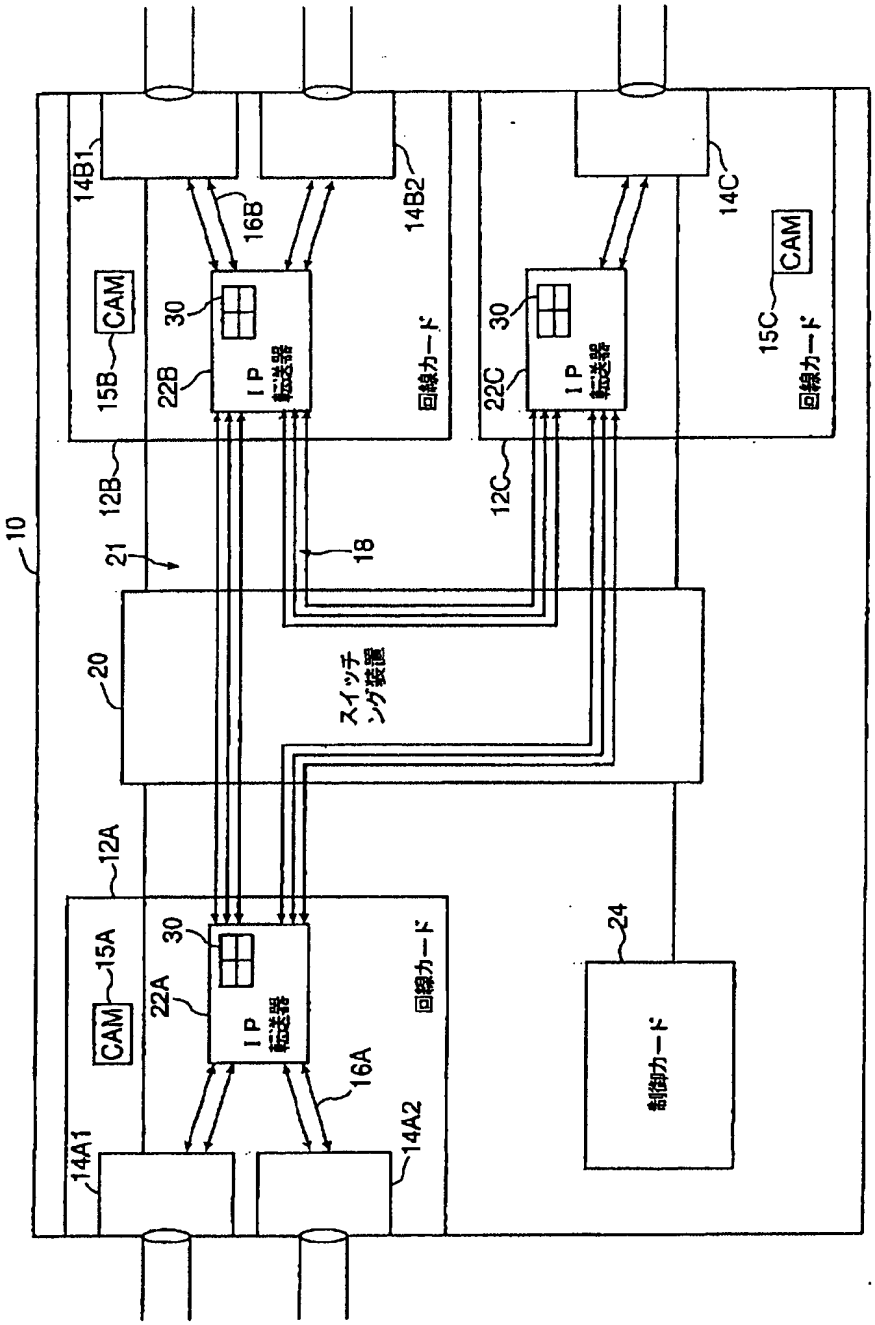

Figure 1

【図２】

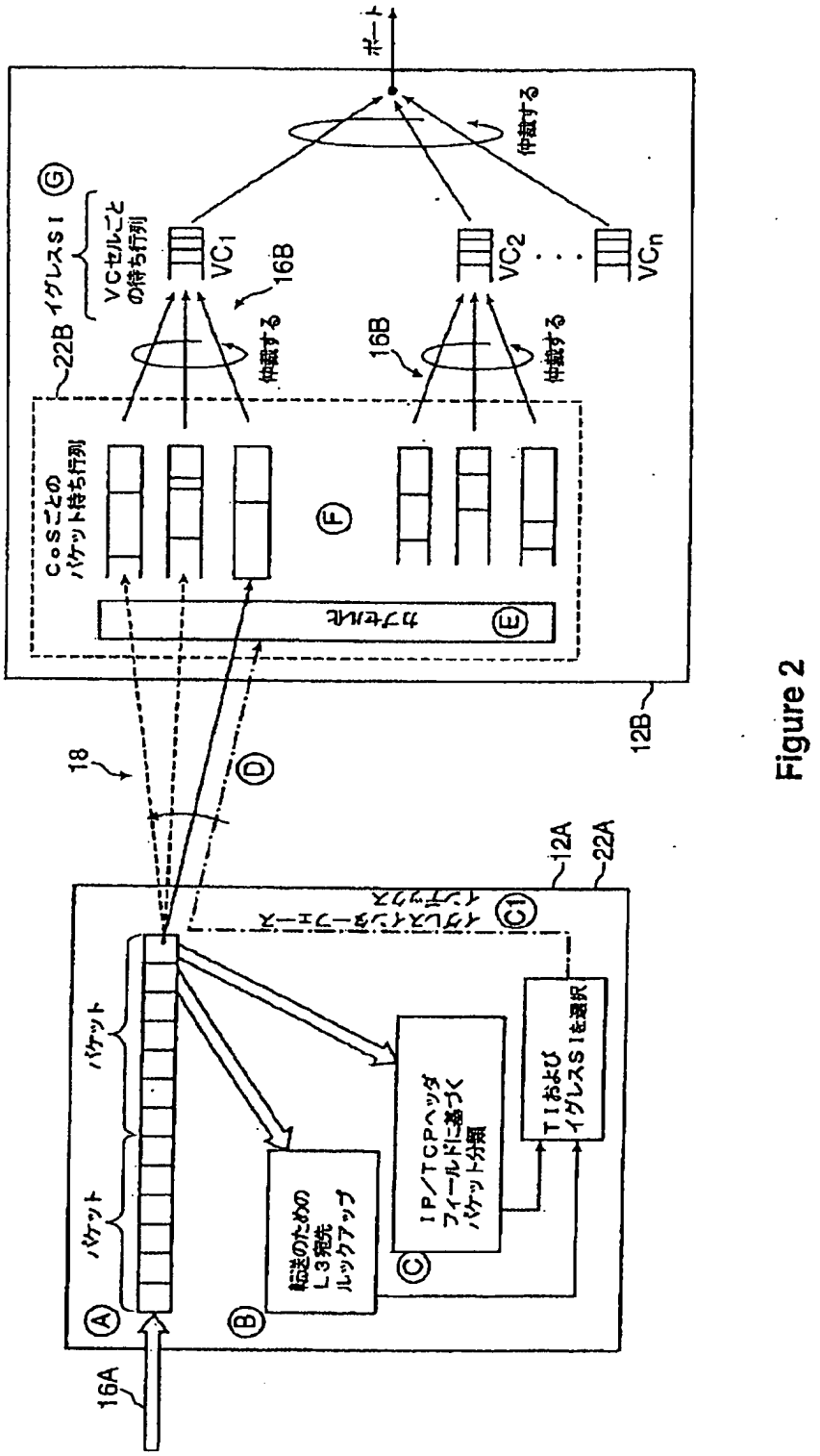

Figure 2

【図３】

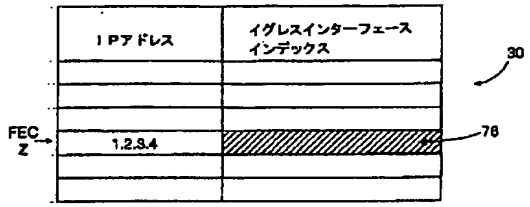

Figure 3

【図６】

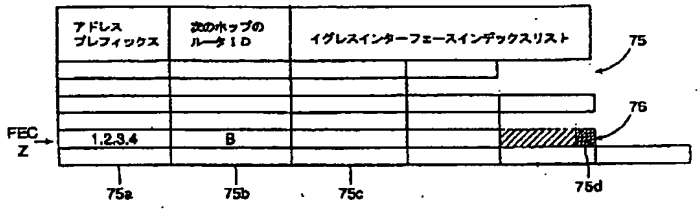

Figure 6

【図７】

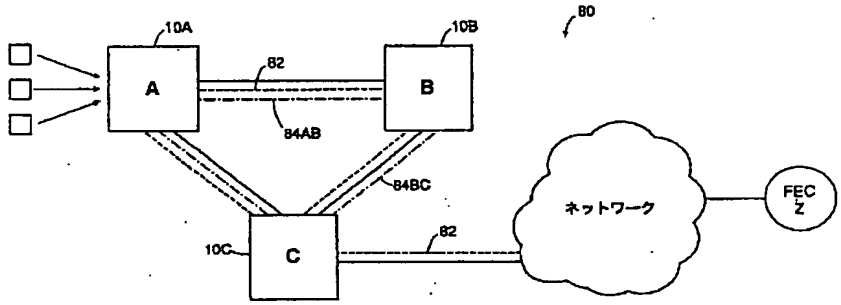

Figure 7

【図８Ｂ】

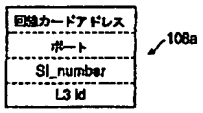

Figure 8B

【図８Ｃ】

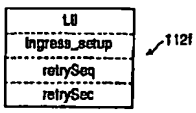

Figure 8C

【図４】

Fig. 4

パラメータ	説明	デフォルト	NMTI	SNMP	NCI	CLI
ＩＤ番号	このＳＩに割り当てられた識別番号；サブスロット内で固有。この値は、内部で割り当てられ、変更することができない。これは5桁の数のフィールドである。	なし	R	-	-	-
端点	ＳＩによって使用されるＡＴＭ端点（シェルフースロットーサブスロットーポート：ＶＰＩ／ＶＣＩ）。	なし	R/W	-	-	-
名前	ＳＩの名前。これは１６文字テキストストリングフィールドである。	空	R/W	-	-	-
アプリケーション	このＳＩによって提供されるアプリケーション。これは、転送、経路指定、およびＬＤＰのそれぞれが使用可能にされているかどうかを示すブールベクトル（すなわち、ビットマップ）である。	転送	R/W	-	-	-
アドレス型	ＩＰアドレスフィールドの型。サポートされる有効な型は、無番号型およびＩＰｖ４型である。	無番号	R/W	-	-	-
ＩＰアドレス	サービスインターフェースのＩＰアドレス。ユーザに対して標準の「小数点付き１０進数」で表示される。「不正」ＩＰアドレス（例えば、０．０．０．０．２５５．２５５．２５５．２５５）がブロックされる。	割当てなし	R/W	-	-	-
ＩＰアドレスプレフィックス長	（サブ）ネットワークＩＤを構成するＩＰアドレス内のビット数。０～３２の範囲の数。	なし	R/W	-	-	-
隣接アドレス型	隣接ＩＰアドレスフィールドの型。サポートされる有効な型は、無番号型およびＩＰｖ４型である。	無番号	R/W	-	-	-
隣接ＩＰアドレス	隣接ルータにあるＳＩの終端で使用されるＩＰアドレス。ユーザ対して標準の「小数点付き１０進数」で表示される。「不正」ＩＰアドレス（例えば、０．０．０．０．２５５．２５５．２５５．２５５）がブロックされる。	割当てなし	R/W	-	-	-
カプセル化	ＳＩ上で使用されるカプセル化。（ＲＦＣ１４８３　ＬＬＣ／ＳＮＡＰ経路指定されたＩＰ、ＲＦＣ１４８３　ＮＵＬＬ）	RFC1483 NULL	R/W	-	-	-
MTU	最大伝送単位	２０１６オクテット	R	-	-	-
イングレストラフィック契約	イングレストラフィック契約構造は、処置（使用禁止、タグ、廃棄）、コミット済み情報速度（ビット／秒での）、およびバーストサイズ（バイトでの）から構成される。各ＳＩ内に８つのイングレストラフィック契約構造が含まれる。それぞれがＣｏＳに適用される。	使用禁止 ＣＩＲ　０ ＢＳ　０	R/W	-	-	-
ステータス	サービスインターフェースのステータス。（アップ、ダウン）。	ダウン	R	-	-	-

サービスインターフェースパラメータ

【図５】

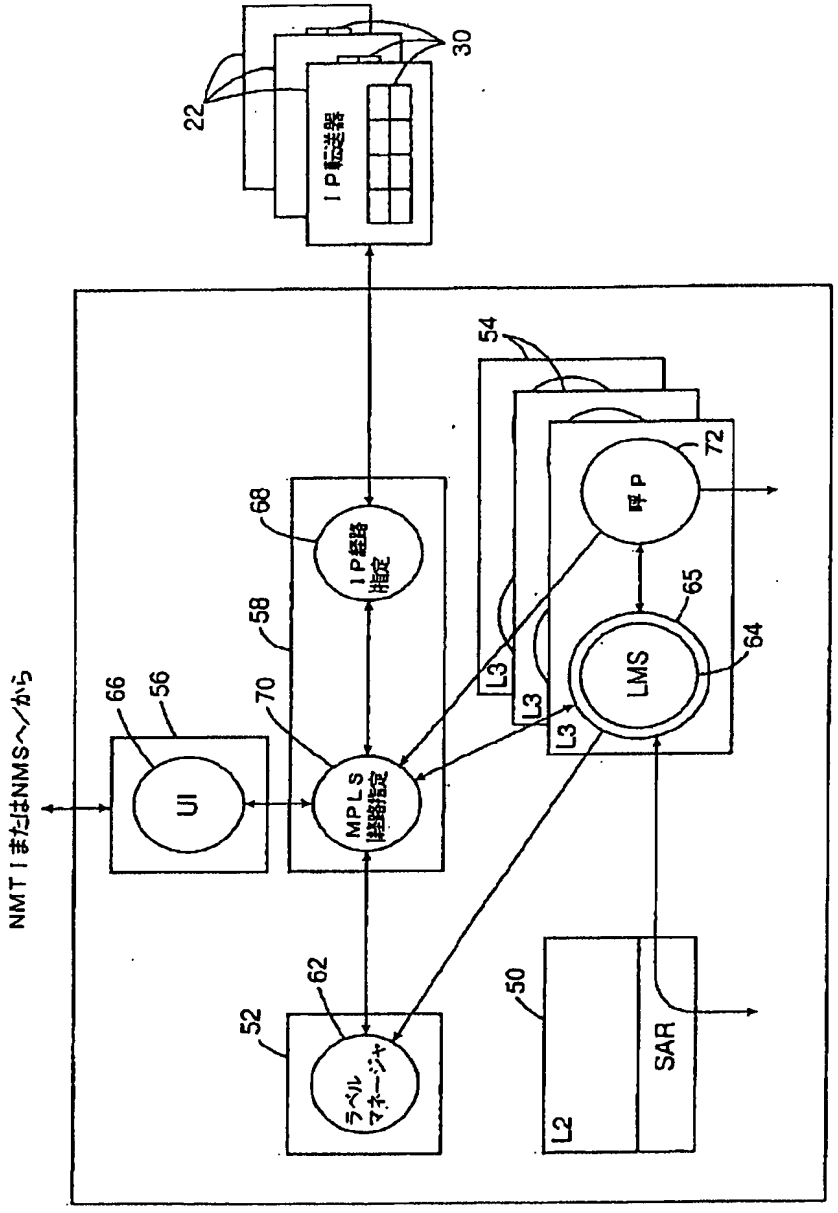

Figure 5

【図８Ａ】

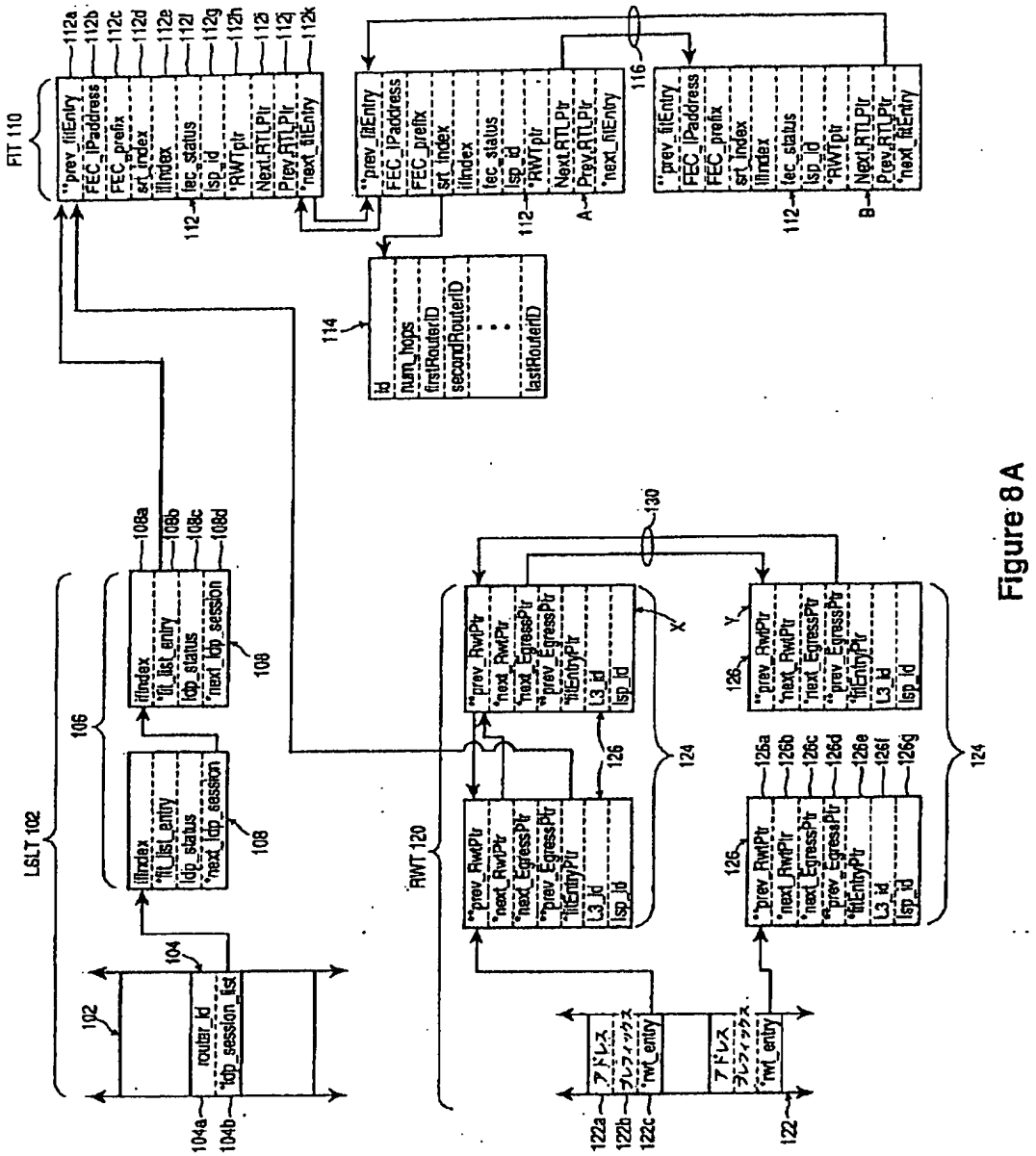

Figure 8A

【図９】

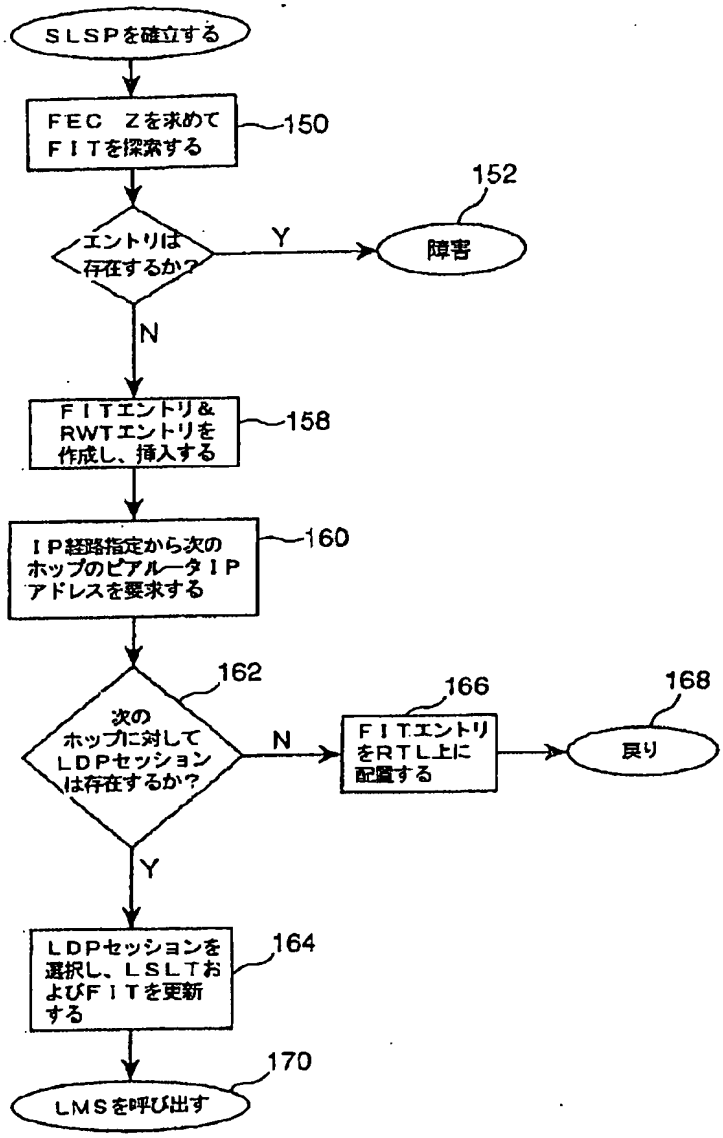

Figure 9

【図１０】

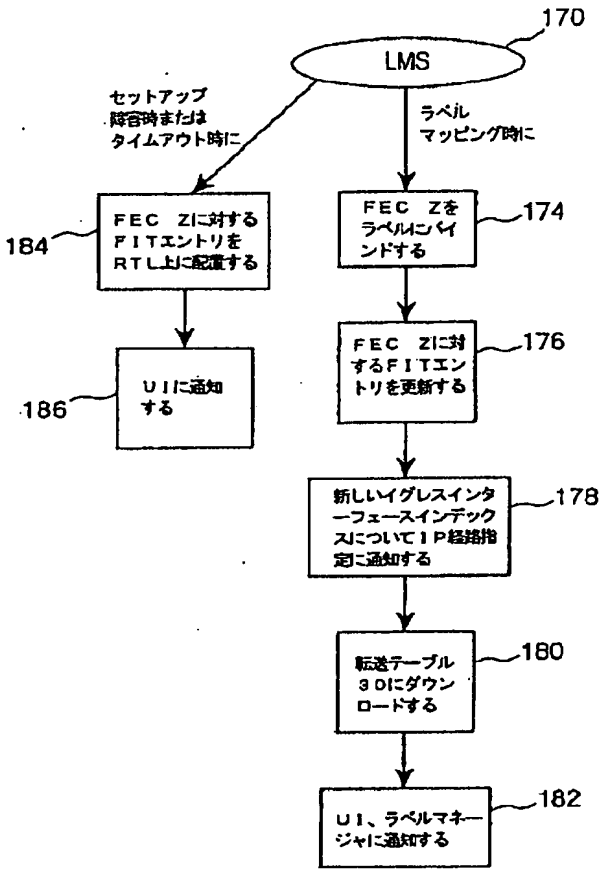

Figure 10

【図１１】

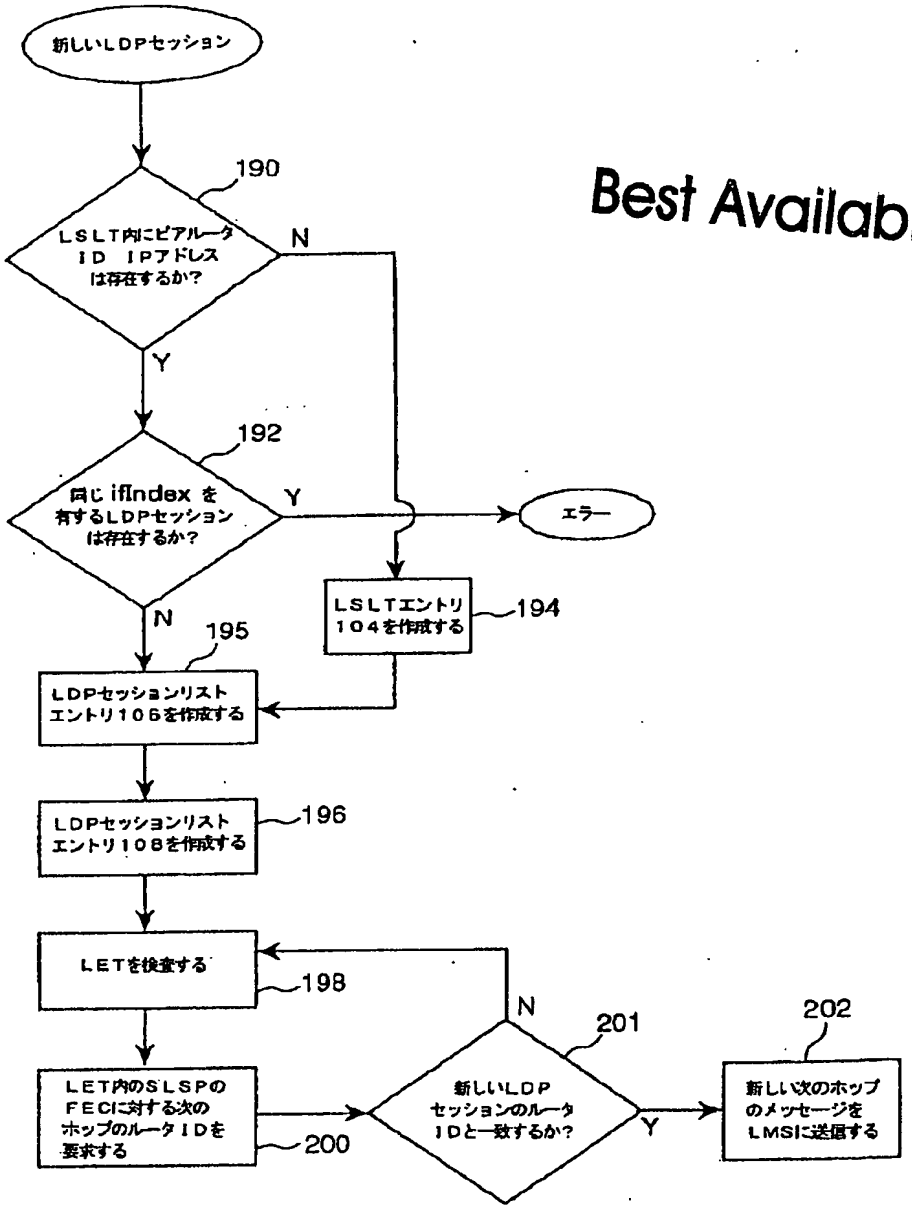

Figure 11

フロントページの続き

(72)発明者 ヌータン・ベーキ
カナダ国、オンタリオ・ケー・２・ジー・０・エム・７、ネピーン、ムーアクロフト・ロード・40

(72)発明者 デイビツド・トスカーノ
カナダ国、オンタリオ・ケー・２・ジエイ・４・エヌ・９、ネピーン、バーントウツド・アベニユー・２

(72)発明者 ケン・ドユブク
カナダ国、オンタリオ・ケー・２・ビー・８・ケイ・３、オタワ、コノート・アベニユー・811、ユニツト・91

Ｆターム(参考) 5K030 GA12 HA10 JL07 LB02 LB08

【外国語明細書】

1. Title of Invention

AN MPLS IMPLEMENTATION ON AN ATM PLATFORM

2. Claims

1. A method of timing an attempt to establish a connection path between a first and second node in a communications network, said method comprising initiating said attempt to establish a connection path after a period of time has elapsed wherein said period of time is greater than another period of time which had previously elapsed between two previous attempts, if any, to establish said connection.

2. The method as claimed in claim 1, wherein said period of time is greater than said another period of time by a fixed time value.

3. The method as claimed in claim 1, wherein said period of time does not exceed a maximum time value.

4. The method as claimed in claim 1 wherein said connection path is a soft permanent label switched path.

5. The method as claimed in claim 2 wherein said fixed time value is ten seconds.

6. A method of timing attempts to establish connections for a plurality of requests for connections in a communication network, said method comprising:

having a timer arrangement tracking passage of a regular interval of time;
having a list of records relating said plurality of requests for connections;
selecting one record from said list;
attempting to establish a connection relating to said one record; and
if said connection relating to said one record is established, then
marking said one record as being successful, otherwise, re-attempting to establish said connection at successive intervals increasing by said regular interval.

7. The method as claimed in claim 6 wherein said selecting one record from said list comprises:

having a time field in said list of records;

on each said regular interval of time for each entry in said list of records:

decrementing a time value in said time field; and

if said time value is zero for an entry is zero, then

selecting said entry as said one record.

8. The method as claimed in claim 6, wherein when re-attempting to establish said connection at successive time intervals, said successive time intervals do not exceed a maximum time value.

9. The method as claimed in claim 8 wherein said maximum time value is sixty seconds.

10. In a communications network comprising two nodes having at least two communications links associated between said two nodes, a method of selecting one of said at least two communications links for signalling between said two nodes utilizing a round-robin algorithm.

11. The method as claimed in claim 10 wherein said method further comprises not selecting any communications link of said at least two communications links having insufficient resources for communications between said two nodes or having a failure therein.

3. Detailed Description of Invention

FIELD OF ART

The invention relates to the art of digital communication systems and more specifically to an implementation of a network node employing multi-protocol label switching (MPLS) over an asynchronous transfer mode (ATM) platform.

BACKGROUND OF INVENTION

MPLS is quickly gaining support in the industry as a robust way of transmitting Internet Protocol (IP) packets. This is primarily because MPLS eliminates the need to examine the destination IP address of a packet at every router or network node in the path of the packet. As such, MPLS has particular utility in the high speed core of many networks.. Recognizing that within the high speed core ATM switching infrastructure is likely to exist, the industry is presently in the process of formulating standards for deploying MPLS over an ATM infrastructure.

As in the nature of most standardization efforts, the focus has been to define the functional features necessary to enable interoperability amongst equipment manufactured by a variety of participants. However, many problems arise in implementing MPLS functionality. These include: (1) the general management and maintenance of signalled label switched paths (SLSP) on an ATM infrastructure; (2) procedures on the failed establishment of an SLSP; (3) management of signalling links; (4) procedures when a signalling link or physical link fails; and (5) procedures under various changes in network topology, such as the creation of a new signalling link. The present invention seeks to provide solutions to these various issues.

SUMMARY OF INVENTION

One aspect of the invention provides a method of managing a communications network having a plurality of interconnected nodes wherein a connection path is established from an ingress node to an egress node through a plurality of intermediate nodes. The method includes: associating the connection path with a network-wide unique identification; storing the path identification on the ingress node so

as to indicate that the path originates thereat; storing the path identification on each intermediate node so as to indicate that the path transits each such intermediate node; and storing the path identification on the egress node so as to indicate that the path terminates thereat.

Preferably, the steps of storing the connection identifier occurs in the process of establishing the connection path by signalling a connection set-up request from the ingress node through the intermediate nodes to the egress node.

Another aspect of the invention relates to a method of timing an attempt to establish a connection path, such as an SLSP, which has initially failed. This is accomplished by initiating another attempt to establish a connection path after a period of time has elapsed, wherein said period of time is greater than another period of time which had previously elapsed between two previous attempts, if any, to establish said connection.

Another aspect of the invention relates to method of timing attempts to establish connections for a plurality of requests for connections, such as SLSPS, in a communication network. The method includes: providing a timer arrangement for tracking passage of a regular interval of time; providing a list of records relating to the plurality of requests for connections; selecting one record from the list; attempting to establish a connection relating to the one record; and if the connection relating to the one record is established, marking the one record as being successful, otherwise, re-attempting to establish the connection at successive intervals increasing by the regular interval.

In other aspects, the invention provides various combinations and subsets of the aspects described above.

The foregoing and other aspects of the invention will become more apparent from the following description of specific embodiments thereof and the accompanying drawings which illustrate, by way of example only, the principles of the invention. In the drawings, where like elements feature like reference numerals which may bear unique alphabetical suffixes in order to identify specific instantiations of like elements),

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENT

The description which follows, and the embodiments therein, are provided by way of illustrating an example, or examples, of particular embodiments of principles of the present invention. These examples are provided for the purpose of explanation, and not limitations, of those principles. In the description which follows, like elements are marked throughout the specification and the drawings with the same respective reference numerals.

1. Overview of ATM Switching

Fig. 1 is an architectural block diagram of an exemplary dual function ATM switch and IP router 10 (hereinafter "node"). The node 10 comprises a plurality of input/output controllers such as line cards 12 which have physical interface input/output ports 14. Generally speaking, the line cards 12 receive incoming ATM cells on ports 14. Each ATM cell, in accordance with standardized ATM communication protocols, is of a fixed size and incorporates a virtual path identifier (VPI) and a virtual channel identifier (VCI) so that the cell can be associated with a particular virtual circuit (VC). For each such cell received, the line cards 12 consult a lookup table or content addressable memory (CAM) 15 keyed on VCs. The CAM 15 provides pre-configured addressing information as to the outgoing port and egress line card for each cell. This is accomplished by way of an "egress connection index", which is a pointer to a pre-configured memory location on the egress line card that stores a new VC identifier that should be attributed to the cell as it progresses its way over the next network link. The ingress line card attaches the addressing information and egress connection index to each cell and sends it to a switching fabric 20 which physically redirects or copies the cell to the appropriate egress line card. The egress line card subsequently performs the pre-configured VPI/VCI field replacement and transmits the cell out of the egress port. Further details of this type of ATM switching mechanics can be found in PCT publication no. WO95/30318, all of which is incorporated herein by reference.

The node 10 also features a control card 24 for controlling and configuring various node functions, including routing and signalling functions, as described in much

greater detail below. The line cards 12 may send data received at ports 14 to the control card 24 via the switching fabric 20.

Each line card supports bidirectional traffic flows (i.e., can process incoming and outgoing packets). However for the purposes of description the following discussion assumes that line card 12A and ports 14A1 and 14A2 provide ingress processing and line cards 12B, 12C and ports 14B1, 14B2, 14C1, 14C2 provide egress processing for data traffic flowing from left to right in Fig. 1.

2. Overview of IP Routing

The node of the illustrated embodiment also enables variable length packets of digital data associated with a hierarchically higher communications layer, such as Internet Protocol (IP), to be carried over the ATM transport layer infrastructure. This is made possible by segmenting each variable length packet into a plurality of ATM cells for transport. Certain VCs may thus be dedicated to carrying IP packets, while other VCs may be exclusively associated with native ATM communications.

When a cell arrives at ingress port 14A1 the line card 12A accesses CAM 15A to obtain context information for the VC of the arriving cell, as previously described. The context information may associate the VC with a "service interface". This is an endpoint to a link layer (i.e. "layer 2") path, such as an AAL5 ATM path, through a network. A number of service interfaces (SIs) may exist on each I/O port 14. These service interfaces "terminate" at an IP forwarder 22 on the same line card in the sense that, as subsequently described, the ATM cells constituting an IP packet are reassembled into the packet, following which IP forwarding procedures (as opposed to ATM switching procedures) are followed.

The essence of IP forwarding is that an IP packet received at one SI is re-transmitted at another SI. Referring additionally to the process flow chart shown in Fig.

2, the forwarding process for IP packets can be logically divided into three transport stages, separated by two processing stages, through the node.

The first transport stage, schematically represented by arrows 16A, carries ATM cells associated with an ingress SI from the ingress port 14A1 to the ingress IP forwarder 22A.

The second transport stage carries IP packets from the ingress IP forwarder 22A across the switching fabric 20 to an egress IP forwarder, e.g., forwarder 22B. This second transport stage is implemented via a "connection mesh" 21. Within the connection mesh eight internal connections or transport interfaces (TIs) 18 are set up between each pair of IP forwarders 22 (only three TIs are shown). The TIs are provided so as to enable different levels or classes of service (COS) for IP packets.

The third transport stage, schematically represented by arrows 16B, carries IP packets from the egress IP forwarder 22B to the egress port, e.g. port 14B1, and egress SI.

The first processing stage occurs at the ingress IP forwarder 22A, where the ATM cells associated with an ingress SI are reassembled into IP packets. This is shown as step "A" in Fig. 2. At step "B" the IP forwarder 22A then examines the destination IP address of the packet in order to determine the appropriate egress SI for the "next hop" through the network. This decision is based on an IP forwarding table 30 (derived from IP protocols, as discussed in greater detail below) shown schematically in Fig. 3. Each record of table 30 includes an IP address field 32 and an "egress interface index" field 36. The IP destination address of the packet is looked up in the IP address field 32 to find the longest match thereto (i.e., the table entry which resolves the packet IP address destination as far as possible). The corresponding egress interface index essentially specifies the egress line card 12B, egress IP forwarder 22B, and the egress SI

for the packet (see more particularly the discussion with reference to Fig. 8A). The egress interface index is attached to the IP packet.

In addition, at step "C" the IP forwarder 22A examines the class of service (COS) encapsulated by the packet. Based partly on the encapsulated COS and internal configuration, the IP forwarder 22A selects one of the second-stage TIs 18 which will reach the egress forwarder 22B with a desired class of service. In order to traverse the switching fabric 20, the ingress IP forwarder 22A re-segments the IP packet into ATM cells (shown schematically as step "D") and attaches addressing information to each cell indicating that its destination is the egress IP forwarder 22B.

The second, smaller, processing stage occurs at the egress IP forwarder 22B, where the egress interface index is extracted from the packet and it is modified at step "E" to match the encapsulation associated with the egress SI. Thus, the VPI/VCI associated with the egress SI is attached to the packet. The packet is then delivered to that egress SI (labelled "G") using the third-stage transport 16B corresponding thereto. In this process the packet is segmented once again into ATM cells which are buffered in cell queues associated with the egress SI and/or output port 14B1. A queuing and possible congestion point (labelled "F") also occurs between the second processing and third transport stage — that is, between the egress IP forwarding module 22B and the egress SI (labelled "G").

It will be seen from the foregoing that effecting IP forwarding functionality on an ATM platform is a relatively involved process, requiring that the IP packet be: (a) reassembled at the ingress IP forwarder 22A, (b) subsequently segmented for transport over the switching fabric, (c) re-assembled at the egress forwarder 22B, and (d) subsequently re-segmented for transmission out of the output port. In addition, a non-trivial IP address lookup has to be performed at the ingress forwarder 22A. These steps have to be performed at each network node and hence increase the latency of end-to-end communication.

3. Introduction to MPLS

In order to avoid having to perform the above procedures on each and every packet, the node 10 provides multi-protocol label switching (MPLS) capability. In conventional IP forwarding, routers typically consider two packets to be in the same "forward equivalency class" (FEC) if there is some address prefix in that router's tables which is the longest match for the destination address of each packet. Each router independently re-examines the packet and assigns it to a FEC.. In contrast, in MPLS a packet is assigned to a FEC only once as the packet enters an MPLS domain, and a "label" representing the FEC is attached to the packet. When MPLS is deployed over an ATM infrastructure, the label is a particular VC identifier. At subsequent hops within an MPLS domain the IP packet is no longer examined. Instead, the label provides an index into a table which specifies the next hop, and a new label. Thus, at subsequent hops within the MPLS domain the constituent ATM cells of a packet can be switched using conventional ATM switching techniques. Such paths are known in the art as label switched paths (LSPs), and LSPs may be manually set up as permanent label switched paths (PLSP) by network operators. Alternatively, a label distribution protocol (LDP) may be employed wherein the network automatically sets up the path upon command from the network operator. Such paths are typically referred to in the art as soft-permanent or signalled LSPs (SLSPs). Further details concerning MPLS can be found in the following draft (i.e. work in progress) MPLS standards or proposals, each of which is incorporated herein by reference:

[1] E. Rosen, A. Viswanathan, R. Callon, *Multiprotocol Label Switching Architecture*, draft ietf-mpls-arch-06.txt.

[2] L. Andersson, P. Doolan, N. Feldman, A. Fredette, B. Thomas, *LDP Specification*, draft-ietf-mpls-ldp-06.txt. This LDP is hereinafter referred to as "LDP Protocol".

[3] B. Davie, J. Lawrence, K. McCloghrie, Y. Rekhter, E. Rosen, G. Swallow, P. Doolan, *MPLS Using LDP and ATM VC Switching*, draft-ietf-mpls-atm-02.txt.

[4] B. Jamoussi, Constraint-*Based LSP Setup using LDP*, draft-ietf-mpls-cr-ldp-01.txt.. This LDP is hereinafter referred to as "CRLDP".

[5] E. Braden et al., *Resource Reservation Protocol*, RFC 2205. This LDP is hereinafter referred to as "RSVP".

The node 10 implements MPLS functionality through an SI linkage, as will be better understood by reference to Fig. 4 which shows the SI in the context of a management entity or record. The SI has an internal ID number associated therewith. In addition to representing an ATM link layer endpoint, the SI also represents an IP address for layer 3 functionality, and indicates what type of encapsulation is used for IP purposes. Each SI may also be associated with a number of other attributes and methods. In particular, SIs can be associated with the following methods or applications: (1) IP forwarding, (2) MPLS forwarding, (3) IP routing, and (4) MPLS signalling. In other words, the node 10 can be configured to (1) forward IP data packets to the next hop router via the above described IP forwarding procedures discussed above; (2) forward IP data packets via MPLS forwarding procedures as will be discussed below; (3) process packets carrying messages for IP routing protocols; and (4) process packets carrying messages for MPLS signalling protocols.

4. Overview of MPLS Architecture

Fig. 5 shows the hardware and software architecture of the control card 24 in greater detail. From the hardware perspective, the card 24 employs a distributed computing architecture involving a plurality of discrete physical processors (which are represented by rectangular boxes in the diagram).

Processor 50 handles layer 2 ("L2") ATM adaptation layer packet segmentation and reassembly functions for signalling messages. As mentioned, certain SIs will be associated with various types of routing protocols and upon receipt of a packet associated with one of these SIs the ingress IP forwarder 22A sends the packet (which is re-segmented to traverse the switching fabric 20) to the L2 processor 50. After re-assembly, the L2 processor 50 sends signalling messages associated with IP routing protocols to a software task termed "IP Routing" 68 which executes on a routing

processor 58. (The connection between the L2 processor 50 and IP Routing 68 is not shown). Signalling messages associated with MPLS LDP protocols are sent to a label management system task (LMS) 64 executing on a layer 3 (L3) processor 54. Outgoing messages from the LMS 64 and IP Routing 68 are sent to the L2 processor 50 for subsequent delivery to the appropriate egress line card and egress SI.

IP Routing 68 runs an IP interior gateway or routing protocol such as I-BGP, ISIS, PIM, RIP or OSPF. (The reader is referred to http://www.ietf.org./html.charters/wg-dir.html for further information concerning these protocols.) As a result of these activities IP Routing 68 maintains a master IP routing table 75 schematically shown in Fig. 6. Each record of the master table 75 includes a field 75a for an IP address field, a field 75b for the next hop router ID (which is an IP address in itself) corresponding to the destination IP address or a prefix thereof, and a list 75c of egress interface indexes associated with the next hop router ID. As the network topology changes, IP Routing will update the forwarding tables 30 of IP forwarders 22 by sending the appropriate egress interface index to it. (Note that table 30 only has one egress interface index associated with each IP destination address entry.)

As shown in Fig. 5, the node 10 employs a plurality of L3 processors 54, each of which includes an LMS 64. Each LMS 64 terminates the TCP and UDP sessions for the LDP signaling links (LDP Session) and runs a state machine for each LSP. As discussed in greater detail below, the LMS 64 receives requests to set up and tear down LDP Sessions, and to set up and tear down SLSPs.

The LMS 64 is commercially available from Harris & Jeffries of Dedham, MA. For intercompatibility purposes, the node 10 includes "translation" software, the MPLS context application manager (MPLS CAM) 65, which translates and forwards incoming or outgoing requests/responses between the LMS and the remaining software entities of the control card 24.

Each L3 processor 54 also includes a call-processing task 72. This task maintains state information about connections which have been requested.

Another processor 56 provides user interface functionality, including interpreting and replying to administrative requests presented through a central network management system (NMS) (such as the Newbridge Networks Corporation 46020™ product) or through command instructions provided directly to the node via a network terminal interface (NTI). For MPLS functionality, a user interface 66 is provided for accepting and replying to management requests to program PLSPs, SLSPs, and LDP Sessions.

A resource control processor 52 is provided for allocating and de-allocating resources as connections are established in the node. For MPLS functionality, processor 52 includes a label manager task 62 which allocates unique label values for LSPs.

On the routing processor 58, a software task termed "MPLS Routing" 70 interfaces between the UI 66, IP Routing 68 and the LMSs 64 running on the L3 processors 54. Broadly speaking, MPLS Routing 70 manages SLSPs. For example, during path setup, MPLS Routing 70 receives an SLSP setup request from the user interface 66, retrieves next hop routing information from IP Routing 68, chooses an LDP Session to the next hop, and calls the appropriate instantiation of the LMS 64 to set up the SLSP path using the selected LDP Session. When a label mapping is received for the path, the LMS 64 informs MPLS Routing 70. MPLS Routing 70 then triggers an update to the forwarding tables 30 of the IP forwarders 22 for the new path. Similarly, when the network topology changes, MPLS Routing 70 reflects these changes into the MPLS routing domain. The functions of MPLS Routing are the focus of the remainder of this description.

5. Reference Network

Fig 7 shows a reference IP network 80 wherein an MPLS routing domain exists amongst routers/nodes A, B and C, the remaining of the network 80 employing IP specific routing protocols such as OSPF. Assume the network operator wishes to establish an SLSP, commencing from node A, for IP destination address 1.2.3.4 (hereinafter "FEC Z") located somewhere in the network. (Note that a FEC, as per the draft MPLS standards, comprises a destination IP address and a prefix thereof.) The network operator may enter a management command at node A via its NMTI or the NMS (not shown) requesting the establishment of a SLSP for FEC Z. Depending on the type of label distribution protocol employed (e.g., LDP Protocol, CRLDP, or RSVP) the network operator may specify the destination node for the SLSP, or even explicitly specify the desired route for the SLSP up to some destination node (i.e., a source-routed SLSP). In the further alternative, the label distribution protocol may use a best effort policy (e.g., in LDP Protocol) to identify nodes (within the MPLS routing domain) as close as possible to the destination address of FEC Z. In the illustrated reference network, assume that node C is the "closest" node within the MPLS routing domain for FEC Z.

In the network 80, signalling links 82 (which are associated with particular SI's) are provided for communicating IP routing messages between the nodes. In addition, signalling links 84 are provided for communicating MPLS label distribution protocol messages therebetween. Each signalling link 84 has an LDP Session associated therewith.

For the purposes of nomenclature, unless the context dictates otherwise, the term "ingress SLSP" is used to identify the SLSP at the originating node (e.g., node A), the term "transit SLSP" is used to identify the SLSP at transiting nodes (e.g., node B), and the term "egress SLSP" is used to identify the SLSP at the destination node (e.g., node C).

The reference IP network shown in Fig. 7 is used to provide the reader with a typical application which will help to place the invention in context and aid in explaining it. Accordingly, the invention is not limited by the particular application described herein.

6. Database Management of SLSPs

In order to create, monitor and keep track of ingress, transit and egress SLSPs, MPLS Routing 70 maintains a number of tables or data repositories as shown in the database schema diagram of Fig. 8. Since each SLSP is managed by an LDP Session, MPLS Routing 70 on each node keeps track of the various LDP Sessions which have been set up between the node and its LDP peer routing entities using an LDP signalling database (LSLT) 100. The LSLT 100 comprises a hash table 102 having one entry or record 104 per LDP routing peer. Record 104 contains a router id field 104a which functions as the index for the hash table 102 and a pointer 104b (i.e., *ldp_session_list) which points to an LDP session list 106. The router id field 104a stores the IP address of the LDP peer router to which one or more LDP Sessions have been configured. Each LDP Session is represented by an entry or record 108 in the pointed-to LDP session list 106. Note that multiple LDP Sessions can be configured between the node and a given MPLS peer router and hence the session list 106 can have multiple entries or records 108. In Fig. 8, two LDP Sessions have been configured with respect to the LDP peer router identified by the illustrated router id field 104a, and hence two records 108 exist in the corresponding LDP session list 106.

Each record 108 of the LDP session list 106 comprises the following fields:

- ifIndex (108a) - A unique number within the node 10 which identifies a particular interface index and SI which has been configured for the LDP application. Fig. 8A shows the structure of the ifIndex field in greater detail. It comprises a node-internal device address for the line card/IP module responsible for the SI, the egress port, the SI ID number (which is only unique per line card) and an identification code or

internal device address for the L3 processor 54 on the control card 24 handling the LDP signalling link.

- *fit_list_entry (108b) - A pointer to a FEC information table (FIT) 110. The FIT, as described in greater detail below, keeps track of all ingress SLSPs stemming from the node. The fit_list_entry pointer 108b points to a list within FIT 110 of the ingress SLSPs associated with this LDP Session.

- ldp_status (108c) - A status indication. The status includes a one bit flag (not shown) indicating whether or not the LDP Session is in use and a one bit flag (not shown) indicating whether resources are available for the LDP Session. An LDP Session is considered to have no resources available when there are no labels available for allocation or when the associated SI becomes non-operational.

- *next_ldp_session - A pointer to another LDP Session record 108 associated with the same LDP peer router.

The FIT 110 keeps track of ingress SLSPs, i.e., SLPs which have commenced from the node. (Note that the FIT 110 does not keep track of transit or egress SLSPs) A FIT entry or record 112 is created by MPLS Routing 70 when an SLSP is configured and record 112 is removed from the FIT 100 when an SLSP is deleted.

Each FIT entry or record 112 comprises the following elements:

- **prev_fitEntry (112a) – A pointer to a pointer which references the current entry. This is used for ease of addition and removal from a list.

- FEC – IP destination for an LSP. The FEC consists of an IP destination address 112b and a prefix 112c for which the LSP is destined, as per the draft standards.

- Srt_index (112d)- An index into a source-route table or list (SRT) 114. This takes on the value 0 if the LSP is not source-routed and >0 if it is. In the event the SLSP establishment command includes a source routed path, the router ID IP addresses are stored in the SRT 114 in sequential order, as shown.

- ifIndex (112e) – Specifies the egress line card and egress SI used to reach the next hop router for the FEC. The structure of this field is the same as shown in Fig. 8A. Note, however, that in the FIT 110 this field 112e specifies the SI for the egress data path (as opposed to signaling channel) for the FEC.

- fecStatus (112f) – The state of this FIT entry as represented (see Fig. 8B) by a ttl value, an ingressSetup flag, a retrySeq counter, and a rertySec counter. The ttl value indicates a time to live value that should be decremented from the incoming packets. The ingresSetup flag indicates that the SLSP is successfully established. The retrySeq counter keeps track of the number of times MPLS Routing has tried to set up this SLSP, as described in greater detail below. The retrySec counter keeps track of how many seconds are left until the next retry is attempted.

- lsp_id (112g) – A unique identifier used to identify an SLSP within an MPLS domain. In the present embodiment the identifier comprises a concatenation of the node's IP router ID plus a unique number selected by the UI 66 to uniquely identify the LSP within the node. The lsp_id is also used as a hash key for the FIT 110.

- *RWTptr (112h) – A pointer to a route watch database (RWT) 120 described in greater detail below.

- Next.RTLPtr (112i), prev.RTLPtr(112j) – Forward and backward pointers used to keep track of FIT entries 112 in which the ingressSetup flag of the fecStatus field 112f indicates that the corresponding SLSP has not been successfully set up. These pointers are basically used to implement a retry list (RTL) 116 which is embedded in

the FIT 110. For example, the FIT entries 112 labelled "A" and "B" form part of the RTL 116. The RTL thus enables the node to quickly traverse the FIT 110 to find pending SLSPs for all peer routers.

- *next_fitEntry (112k) – A pointer to the next FEC/FIT entry which has been set up using the same LDP Session as the current FEC/ FIT entry.

The RWT 120 keeps track of all SLSPs handled by the node, i.e., ingress, transit and egress SLSPs. The RWT 120 comprises a hash table 122 which includes an IP designation address field 122a, an IP prefix field 122b, and a *rwt-entry 122C which points to a list 124 of LSPs described in greater detail below.

The IP destination address and prefix fields 122a and 122b are used to store different types of management entities depending on the particular label distribution protocol employed. These entities may be: (a) the FEC, for LDP Protocol; (b) the destination node's router ID, for non source-routed RSVP; (c) the next node's router ID for strict source-routed CR-LDP and RSVP; and (d) the next hop in the configured source-route for loose source-routed CR-LDP and RSVP. These can all be summarized as the next hop that an SLSP takes through the network.

Note that table 122 is hashed based on the IP prefix field 122b. There can be several requested SLSPs all referring to the same IP prefix at a transit node or egress node. Each individual SLSP is identified by a separate entry or record 126 in the LSP list 124. However, there can only be one ingress SLSP associated with any given IP prefix on the node 10. (In other words, an entry 126 exists for every next hop request received from the LMS 64 as well as one entry for an ingress SLSP which has been created on the node. Note too that egress SLSPs also request next hop information and therefore are included within this table).

Each LSP list entry 126 comprises the following elements:

- prev_RwtPtr (126a), next_RwtPtr (126f) – Forward and backward pointers used to keep track of additional entries 126 for a specific IP prefix. All of the LSPs associated with the same IP prefix 122b are linked together using pointers 126a and 126f.

- next_EgressPtr (126b), prev_EgressPtr (126c)– Forward and backward pointers used to keep track of egress SLSPs which may possibly be extended when a new LDP Session is configured, as discussed in greater detail below. These pointers are basically used to implement an LSP egress table or list (LET) 130 which is embedded in the RWT 120. For example, in Fig. 8 the RWT entries 126 labelled "X" and "Y" belong to the LET 130. An entry 126 is added to the LET 130 whenever a best effort routing policy (e.g., LDP Protocol) is employed in setting up an SLSP and the node 10 can find no further LDP signalling links "closer" to the destination address of the corresponding FEC. For example, in establishing an SLSP for FEC Z in the reference network, node C (which lies at the boundary of the MPLS routing domain) cannot find any more LDP signalling links heading towards the destination address of FEC Z, and thus when node C creates a RWT entry 126 for this SLSP the entry will be added to the LET.

- fitEntryPtr (126d) – Pointer to the FIT entry 112 which corresponds to this RWT entry 126. The value of this field will be null for all entries except for ingress SLSPs created at this node.

- L3_id (126e) – The address or identity of the L3 processor which initially requested the next hop request for the LSP or the address or identity of the L3 processor which is used to set up an ingress SLSP.

- lsp_id (126g) - Same as lsp_id 112g in FIT 110, except that these LSPs may have been initiated at other nodes.

7. Establishing an LDP Session

LDP Sessions are configured via management requests which are received through the UI 66 and forwarded to MPLS Routing 70. The data obtained by the UI 66 includes the ATM link layer end point of the LDP signalling link SI (i.e. – line card address, port, VPI/VCI), IP address assigned to the SI, and LDP specific parameters such as label range, label space ID and keep-alive timeout.

MPLS Routing 70 employs a round-robin algorithm to select one instantiation of the LMS 64 (i.e., one of the L3 processors 54) and requests the associated MPLS CAM 65 to establish a new LDP Session. The MPLS CAM enables the LDP signalling application on the SI selected by the network operator and configures the node, including a filtering mechanism (not shown) associated with the L2 processor 50, to allow all LDP packets associated with a particular LDP signalling SI to be propagated (in both the ingress and egress directions) between the line cards 12 and the selected LMS/L3 processor 54. Once this is carried out, the LMS 64 sends out LDP session establishment messages to the LDP peer router in accordance with the applicable label distribution protocol (e.g., LDP Protocol, CRLDP, RSVP). These include "hello" and other session establishment messages.

Once an LDP Session has been established with the LDP peer router, the LMS 64 informs the label manager 62 of the negotiated label range for the LDP Session (which is a function of establishing an LDP Session as per the draft standards). The LMS 64 also passes the IP address of the LDP peer router to MPLS Routing 70 which stores this address in the router ID field 104a of the LSLT100. In addition, the LMS 64 passes the interface index identifying the LDP signalling SI to MPLS Routing 70 which stores it in the ifIndex field 108a of the LSLT 100.

8. Establishing an SLSP

8.1 *Procedures at the Ingress Node*

Referring to the reference network, an SLSP must be explicitly established at node A for FEC Z by the network operator via the NMTI of node A or via the NMS which communicates with node A. The instruction to configure the SLSP includes as one of its parameters Z, i.e., the destination IP address and prefix thereof for FEC Z. The command is received and interpreted by the UI 66.

The UI 66 selects a unique LSP ID which, as previously discussed, preferably comprises a concatenation of the node's IP router ID and a unique number. The UI 66 then requests MPLS Routing 70 to create an SLSP for FEC Z and associate it with the selected LSP ID.

MPLS Routing 70 requests next hop information for FEC Z from IP Routing 68. This will occur for non-source-routed LSPs in order to obtain the next-hop information as well as for source-routed LSPs in order to verify the information in the source-route (which will be supplied by the network operator). More specifically, MPLS Routing 70 executes the following procedure to initiate the establishment of an SLSP for this new FEC.

Referring additionally to Fig. 9, at a first step 150 MPLS Routing 70 searches the FIT 110 for an existing entry 112 having the same IP destination address and prefix as FEC Z. If such an entry exists in the FIT 110 then at step 152 MPLS Routing 70 returns with a failure code indicating that FEC Z has already been established from this node. At step 158, MPLS Routing 70 creates a new FIT entry 112 and appends it to the FIT 110. A corresponding entry 126 is also inserted into the LSP list 124 for FEC Z in the RWT hash table 122. If necessary, MPLS Routing 70 adds a new entry 122 to the RWT 120 which includes the IP prefix and address of FEC Z, or the IP prefix and address of the first hop in the explicit route.

At step 160 MPLS Routing 70 requests IP Routing 68 to provide the peer IP address for the next hop to reach FEC Z (or the destination node's router id for non source-routed RSVP, or the next hop in the configured source-route for loose source-routed CR-LDP and RSVP). Once obtained, at step 162 MPLS Routing 70 searches for an LSLT entry 102 which matches the next hop router ID. If a matching LSLT entry exists, then at step 164 MPLS Routing 70 selects an available LDP Session from the corresponding LDP Session list 106. This is a circular linked list, which is managed such that the *ldp_session_list pointer 104b in the LSLT entry 102 points to the LDP Session to be used for the next SLSP setup which is selected by MPLS Routing 70. Once the LDP Session is selected, the recently created FIT entry 112 for FEC Z is linked (via the **prev_fitEntry and *next-FitEntry pointers 112a and 112i) to other FIT entries using the same LDP Session.

The *next_ldp_session pointer 108d points to the next session in the LDP session list. (If there is only one LDP Session in the list then the *next_ldp_session points to itself.) Once the link between the FIT 110 and LDP session list 106 is created, MPLS Routing 70 updates the *ldp_session_list pointer 104b to point to the next session in the LDP session list with resources. This results in a round robin approach to selecting LDP Sessions for a given FEC. If no sessions to the peer LDP router have resources, the ldp_session_list pointer 104b is not updated. In this case, the list is traversed once when a path is setup before MPLS Routing 70 stops looking for a session.

Note also that if MPLS Routing 70 does not find an LSLT entry 102 which matches the next hop router ID, then no LDP signaling link exists thereto. In this case MPLS Routing 70 adds the recently created FIT entry for FEC Z to the RTL at step 166 and returns at step 168 with an appropriate failure code.

Once an LDP Session has been selected to signal the establishment of the SLSP, then at step 170 MPLS Routing 70 requests the LMS 64 to signal the set up of an SLSP. The LMS 64 of node A sends a label request message, as per the draft LDP

standards, to its downstream LDP peer router, node B, indicating the desire to set up an LSP for FEC Z. The label request message is propagated downstream across the MPLS routing domain in accordance with the routing protocol (hop-by-hop or source routed) to the egress node C, and label mapping messages are propagated upstream back to the ingress node A. Ultimately, as shown in Fig. 10, a label message should be received inbound on the LDP signalling link selected by MPLS Routing 70 for FEC Z. This label message identifies the label, i.e., VPI/VCI value, that should be used to forward IP packets and the ATM cells thereof to node B. The label is passed to MPLS Routing 70 and to the label manager 62. In addition, at step 174 the LMS 64 signals the call processor 72 to configure an egress interface index for the SI being used on the egress line card and port to handle the data traffic. (Note that the egress line card will be the same line card and port associated with the LDP signaling SI for FEC Z.) This "binds" FEC Z to the ATM VPI/VCI label. The binding is reported to MPLS Routing 70 which searches the FIT 110 at step 176 for the entry 112 matching FEC Z, whereupon the ifIndex field 112e is updated with the egress interface index obtained from the call processor 72.

In addition, MPLS Routing 70 updates the fecStatus field 112f (Fig. 8B) by setting the retrySeq and retrySec counters to zero and sets the ingressSetup flag to one thereby indicating successful set up. At step 178 MPLS Routing 70 informs IP Routing 68 about the newly established SLSP and its egress interface index whereupon the latter task updates its IP forwarding table 75 (Fig. 6) to add the newly established egress interface index (shown schematically by ref. no. 76) to the appropriate list 75c. IP Routing 68, in turn, may have a number of potential egress interface indexes in list 75c, which may be used to forward a packet. In order to decide amongst these alternatives, IP Routing 68 employs a priority scheme which grants an MPLS-enabled egress interface index (there can only be one per FEC) higher priority than non-MPLS egress interfaces. The priority scheme is carried out through the mechanism of a bit map 75d (only one shown) which is associated with each entry of the egress interface index list 75c. The bit map 75c indicates what type of application, e.g., SLSP or IP, is associated with the egress interface index entry. Following this priority scheme, at step 180 IP Routing downloads

the newly created egress interface index 76 to the forwarding tables 30 of each IP forwarding module. (Table 30 only lists a single egress interface index for each IP address or prefix thereof). Asynchronously. MPLS Routing 70 also informs the UI 66 at step 182 that the ingress SLSP for FEC Z has been successfully created.

In the event that no label mapping message is received within a predetermined time period, or the signalling message that is received from node B denies the setup of an SLSP for FEC Z, the LMS 64 informs MPLS Routing 70 of the failure at step 184. MPLS Routing consequently places the FIT entry 112 for FEC Z on the RTL 116, sets the fecStatus ingressSetup field (Fig. 8B) to zero and increments the value of the retrySeq field (up to a max of 6). At step 186, MPLS Routing informs the UI 66 of the failure.

The retry mechanism for FIT entries is a linear back off mechanism which causes an SLSP path setup to be retried at 10, 20, 30, 40, 50, and 60 seconds. There is one retry timer associated with MPLS Routing 70 which goes off every 10 seconds. At this point MPLS Routing traverses the RTL 116, decrementing the amount of time (retrySec – Fig.8B) left for each FIT entry 112 in the RTL 116. If the retrySec value is zero, the FIT entry 112 is removed from the RTL 116, the retry sequence number is incremented by one and another attempt is made to establish the ingress SLSP. If the retry is successful retrySeq is set to zero and the ingressSetup flag is set to 1. If the retry is unsuccessful then the FIT entry is added back to the RTL, retrySeq is incremented (max. sequence number is preferably 6). When the retrySeq counter is increased, the time period within which MPLS Routing 70 will retry to set up the SLSP also increases to the next highest interval. For instance, when retrySeq increases from 2 to 3 the time interval between retries increases from 20 to 30 seconds, i.e. retrySec is set to 30. When retrySeq is equal to 6, retries are 60 seconds apart.

8.2 Procedures at Transit Nodes

At transit node B, a label request message for FEC Z is received on MPLS signaling link 84 and forwarded by the L2 processor 50 to the responsible LMS 64. The LMS 64 requests next hop information from MPLS Routing 70 which, in turn, retrieves the next hop router ID for FEC Z from IP Routing 68, stores the next hop router ID in the RWT 120, selects a downstream LDP Session to the next hop LDP peer router, node C, and supplies this data to the LMS 64, as discussed previously. The LMS 64 then requests the label manager 62 to reserve a VPI/VCI label from within the negotiated label range (determined when the LDP Session with the upstream node A was established). This label is forwarded upstream to node A when the label mapping message is sent thereto Then, if necessary, the LMS 64 which received the upstream label request message will signal another instantiation of the LMS (on a different L3 processor 54) responsible for the downstream LDP Session in order to progress the Label Request message to node C.

When a label mapping message is received from the downstream signalling link, the LMS 64 signals the call processor 72 to establish a cross-connect between the label, i.e., VPI/VCI, associated with upstream node A and the label, i.e., VPI/VCI, associated with the downstream node C to thereby establish downstream data flow. On the transit node this results in an ATM style cross-connect, as discussed above. In addition, the LMS 64 responsible for the upstream LDP Session to node A forwards a label mapping message to it with the label previously reserved by the label manager 62.

Note that for source-routed SLSPs it may not be necessary for the transit node B to obtain next hop information from IP Routing 70. This is, however, a preferred feature which enables the transit node to confirm through its internal routing tables that the next hop provided in the source route list is accurate (e.g., by checking whether the next hop is listed under the requested IP destination address or prefix). If the explicitly routed next hop cannot be confirmed, then an error can be declared.

8.3 *Procedures on Egress Node*

On the egress node C, a label request message is received on the upstream signalling link with node B and forwarded by the L2 processor 50 to the responsible LMS 64. The LMS 64 requests next hop information from MPLS Routing 70 which, in turn, requests next hop information from IP Routing 68. In this case, however, one of the following circumstances arises: (1) the next hop router ID returned by IP Routing 68 is the current node; or (2) the next hop is found, but no LDP Session exists to the next hop (i.e., the edge of the MPLS domain is reached). In either of these cases, MPLS Routing 70 informs the LMS 64 that the SLSP for FEC Z must egress at this node, whereby the LMS 64 sends a label mapping message to the upstream node B as previously described but does not (and cannot) progress the label request message for FEC Z forward. In this case, MPLS Routing 70 adds an entry 126 in the RWT 120, as previously discussed, but also adds the newly created RWT entry 126 to the LET 130.

In this case, the LMS 64 instructs the call processor 72 to establish an SI configured for IP forwarding. This SI has an ATM endpoint (i.e., VPI/VCI) equal to the VPI/VCI used as the MPLS label between nodes B and C for the SLSP.

9. Switching/Routing Activity

Having described the set up of an SLSP for FEC Z, the manner in which IP packets associated with FEC Z are processed is now briefly described. At the ingress node A the IP packets arrive at port 14A1 in the form of plural ATM cells which the IP forwarder 22A reassembles into constituent IP packets. Once the destination IP address of the received packet is known, the IP forwarder 22A examines its forwarding table 30 for the "closest" entry. This will be the entry for FEC Z that was downloaded by IP Routing 68 in connection with the establishment of the SLSP for FEC Z. Thus, the forwarding table 30 provides the egress interface index 76, comprising the identity or address of the egress line card 12B, egress port 14B1 and egress SI number. The egress interface index is attached to the packet. The ingress IP forwarder 22A also selects a TI 18 to transport the packet over the switching fabric 20 to the egress IP forwarder 22B, based in part on the COS field encapsulated in the packet. The packet is then re-segment for transport across the switching fabric 20 on the selected TI 18 and received by the

egress IP forwarder 22B. The egress IP forwarder 22B, in turn, extracts the egress SI and COS information attached to the packet and modifies it to match the encapsulation indicated by the egress interface index (i.e., egress SI). This includes attaching the VPI/VCI label to the packet. The packet is subsequently segmented into constituent ATM cells and transmitted out of the egress port 14B1 with the VPI/VCI values indicated by the egress SI.

On the transit node B, the ATM cells corresponding to the IP packets are received by an ingress port. The CAM 15 returns an ATM egress connection index, whereby the cells are processed as ATM cells. The ingress line card 12A also attaches internal addressing information retrieved from the CAM 15A to each cell, thereby enabling the cells to be routed to the egress line card which replaces the VPI/VCI value of the cells. The egress line card then transmits the cell using the new VPI/VCI value. Note that in this case the IP forwarding modules 22 were not involved in the switching activity and there was no need to re-assemble and re-segment the IP packet, or perform the IP routing lookup.

On the egress node C, the ATM cells corresponding to the IP packets are received by an ingress port and processed in accordance with the SI configured for the VPI/VCI carried by the cells. This SI is configured such that the cells are sent to the IP forwarding module 22A for re-assembly into the higher-layer IP packets and thereafter processed as regular IP packets.

10. Network Topology Changes

10.1 *New LDP Session*

When a new LDP Session is established on node 10, the LMS 64 signals MPLS Routing 70 about this event and informs it about the interface index for the new LDP Session. The signal arises whether the node is the initiator of the new LDP Session or the respondent. Referring additionally to the flow chart of Fig. 11, at step 190 MPLS Routing searches for the peer router ID IP address in the LSLT 100. If an LSLT entry 194 for this router is found, then at step 192 MPLS Routing 70 examines the

corresponding LDP Session list 106 to ensure that no entries 108 exist for an LDP Session having the same interface index as the new LDP session. If no such entry is found, a new entry 108 is created at step 195. If such an entry is found, an error is returned. If no LSLT entry 104 is found which matches the peer router ID for the newly configured LDP Session, then at step 194 MPLS Routing creates and inserts a new LSLT entry 104, following which the LDP session list entry 106 is created at step 195.

At step 196, MPLS Routing 70 traverses the LET 130. For each RWT entry 126 belonging to the LET, the corresponding FEC is determined from hash table 122, and at step 200 the next hop router ID for that FEC is requested from IP Routing 68. At step 201 the next hop router ID is compared against the peer router ID of the newly configured LDP Session. If no match is found, control returns to step 198, and if a match is found, control passes to step 202. At step 202, MPLS Routing 70 instructs the LMS 64 to send a label request message to the newly reachable peer router for the identified FEC.

10.2 *Signaling Link Failure*

When an LDP Session fails on a node it stops forwarding all SLSPs using the associated VPI/VCI range (stored in the label manager 62) and removes the cross-connects from the node. The node also sends a label withdraw message to the upstream peer for each SLSP associated with the failed LDP Session. For instance, if the MPLS link 84BC (Fig. 7) fails, node B sends a label withdraw regarding FEC Z to the ingress node A. When the label withdraw message is received at the ingress node A, it stops using the path (IP hop-by-hop forwarding is used instead) and immediately re-initiates the steps described previously to re-establish a path for FEC Z. If this does not succeed, then the SLSP for FEC Z is placed on the RTL 116 following which the retry procedures as previously described are effected.

Furthermore, when an LDP Session becomes inoperative in the ingress node A for whatever reason, the LMS 64 informs MPLS Routing 70. As part of this call, the LMS 64 provides MPLS Routing 70 with the peer router ID IP address. MPLS Routing 70 then searches for the peer IP address in the router ID field 104a of the LSLT

100. If there is no entry for the peer IP address, an error is returned. If there is an entry 104a for the peer IP address, the corresponding session list 106 is searched for the failed LDP Session. If there is a matching LDP Session entry 108, it is removed from the session list 106.

The *fit_list_entry pointer 108b of the removed session list entry 106 points to the list of all FIT entries 112 representing all ingress SLSPs using the failed LDP Session. For each of these entries, MPLS Routing 70 immediately tries to re-establish the ingress SLSP as described above to see if there is an alternate LDP Session that may be used to set up the ingress SLSP. If the retry is unsuccessful, the ingress SLSP goes on the RTL 116 and the retry procedures outline above are followed.

10.3 *IP Routing Changes*

Over the course of time, IP Routing 68 may discover a new next hop for FEC Z. For example, in the reference network IP Routing on node B may discover that the next hop for FEC Z should be node D (not shown). Upon such a discovery, IP Routing 68 on node B informs MPLS Routing 70 of the new next hop router ID for FEC Z. MPLS Routing 70 uses the following process to re-route the SLSP for FEC Z: First, it searches for a RWT entry 122 matching the IP prefix address, e.g., FEC Z, which has changed in the IP Routing table 75. In the event no entry 122 is found MPLS Routing returns otherwise it continues and next searches for an LSLT entry 104 that matches the router ID of the new next hop D. If there is an LSLT entry 104 and hence LDP Session to the new router D, MPLS Routing requests the LMS 64 to progress each transit SLSP in the RWT list 124 pointed to by the matching RWT entry 122 using the LDP Session to router D. Thus transit SLSPs are re-routed to the new next hop router D. However, if there is no LSLT entry 102 for the router ID of the new next hop and hence no LDP Session therefor, then MPLS Routing 70 places each transit SLSP in the RWT list 124 corresponding to the old-hop router on the LET 130 and informs the LMS 64 that it should consider such SLSPs as egress SLPs. The LMS 64 , in turn, instructs the call processor 72 to set up egress SIs for the egress SLSPs.

MPLS Routing 70 also searches for a FIT entry 112 which matches the affected FEC. If there is a FIT entry 112 that matches the FEC and the ingress_setup flag of the fec-status field 112f is non zero (i.e., the path is set up), MPLS Routing 70 requests that the LMS 64 close the ingress SLSP by sending a label release message to the downstream routers. MPLS Routing 70 then searches for an LSLT entry 104a that matches the router ID IP address for the new next hop. If there is such an LSLT entry, then an LDP Session is selected from the corresponding LDP session list 106, and the procedures for establishing an ingress SLSP are followed as described above.

10.4 *Physical Link Failure*

When a physical link between two nodes fail, then signaling links 82 and 84 (see Fig. 7) for both MPLS signaling and IP routing fail. In the present embodiment, IP Routing 68 realizes that the link is down and updates its routing table 75 before the LMS 64 realizes that any LDP Sessions thereover are down. This is accomplished by suitably setting "time out" periods for LDP Sessions and signaling sessions in IP Routing such that interface failures are reflected much quicker into IP Routing 68 than MPLS Routing 70. Accordingly, IP Routing 68 informs MPLS Routing 70 about a new next hop router ID for affected SLSPs and as previously described MPLS Routing 70 will reroute these SLSP paths from the current node, using the new next hop router identified by IP Routing 68. This is more efficient than tearing down the affected SLSPs back to the ingress node and resignaling them as would have occurred if MPLS Routing 70 realizes the signaling link is down.

The foregoing embodiment has been described with a certain degree of particularity for the purposes of description. Those skilled in the art will understand that numerous variations and modifications may be made to the embodiments disclosed herein without departing from the spirit and scope of the invention.

4. Brief Description of Drawings

Fig. 1 is a system block diagram of a network node which processes ATM cells and IP packets.

Fig. 2 is process flow diagram showing how IP packets are processed in the node of Fig. 1.

Fig. 3 is a diagram of a forwarding table employed by IP forwarders associated with input /output controllers of the node of Figure 1.

Fig. 4 is a diagram of a data structure representing a "service interface" associated with nodes such as shown in Fig. 1.

Fig. 5 is an architectural block diagram of hardware processors and software processes associated with a control card on the node of Fig. 1.

Fig. 6 is a master IP routing table associated with an IP network.

Fig. 7 is a diagram of a reference network illustrating an MPLS domain within an IP network.

Fig. 8 is a schematic diagram of a database employed by the node of Fig. 1 to manage signalled label switched paths (SLSPs).

Figs. 8A and 8B show certain fields of the database of Fig. 8 in greater detail.

Figs. 9 and 10 are logic flow charts showing the steps executed by the node of Fig. 1 in establishing an SLSP.

Fig. 11 is a logic flow chart showing the steps executed by the node in the event a new SLSP signalling link is established.

Fig. 1

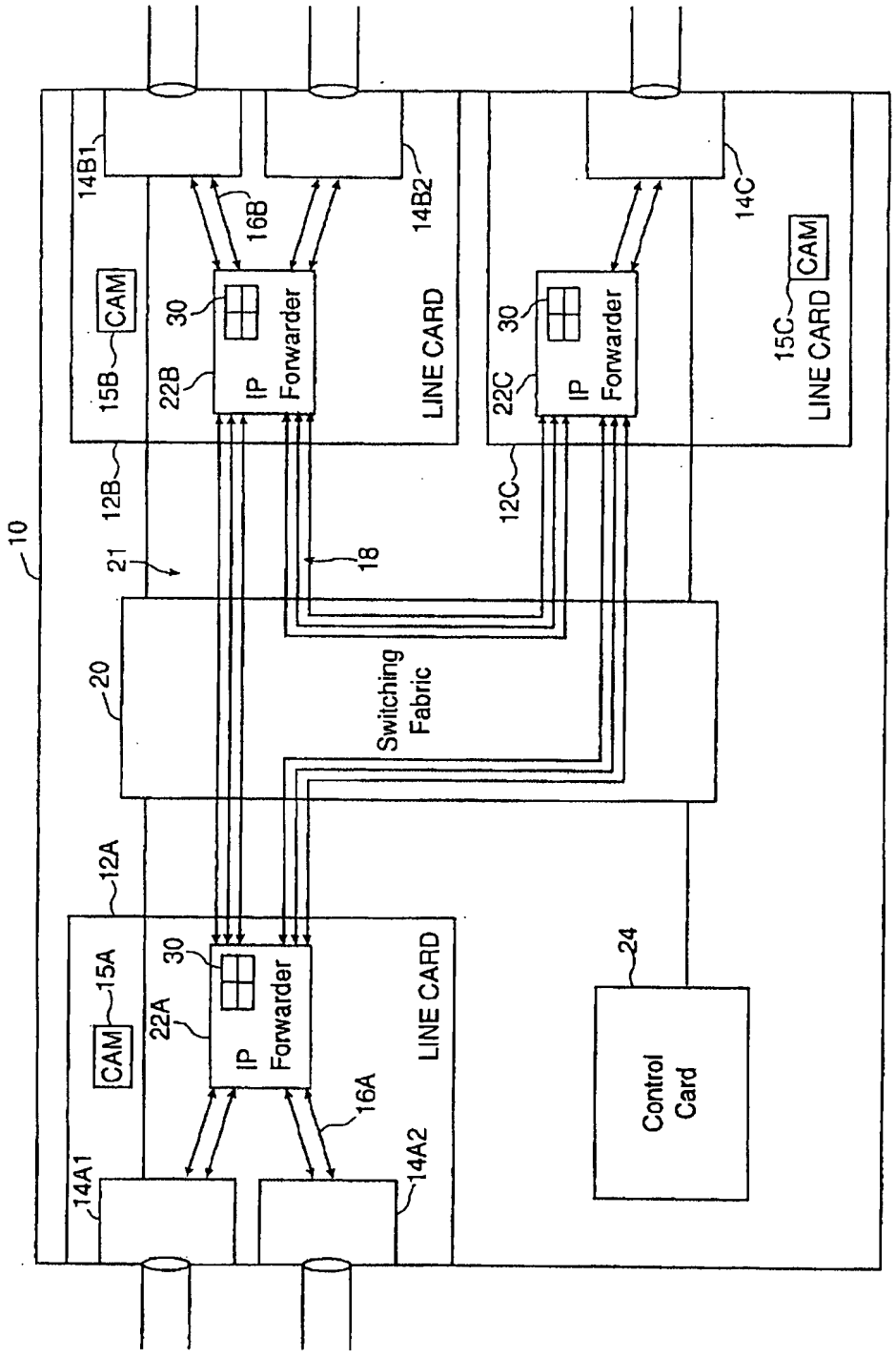

Figure 1

Fig. 2

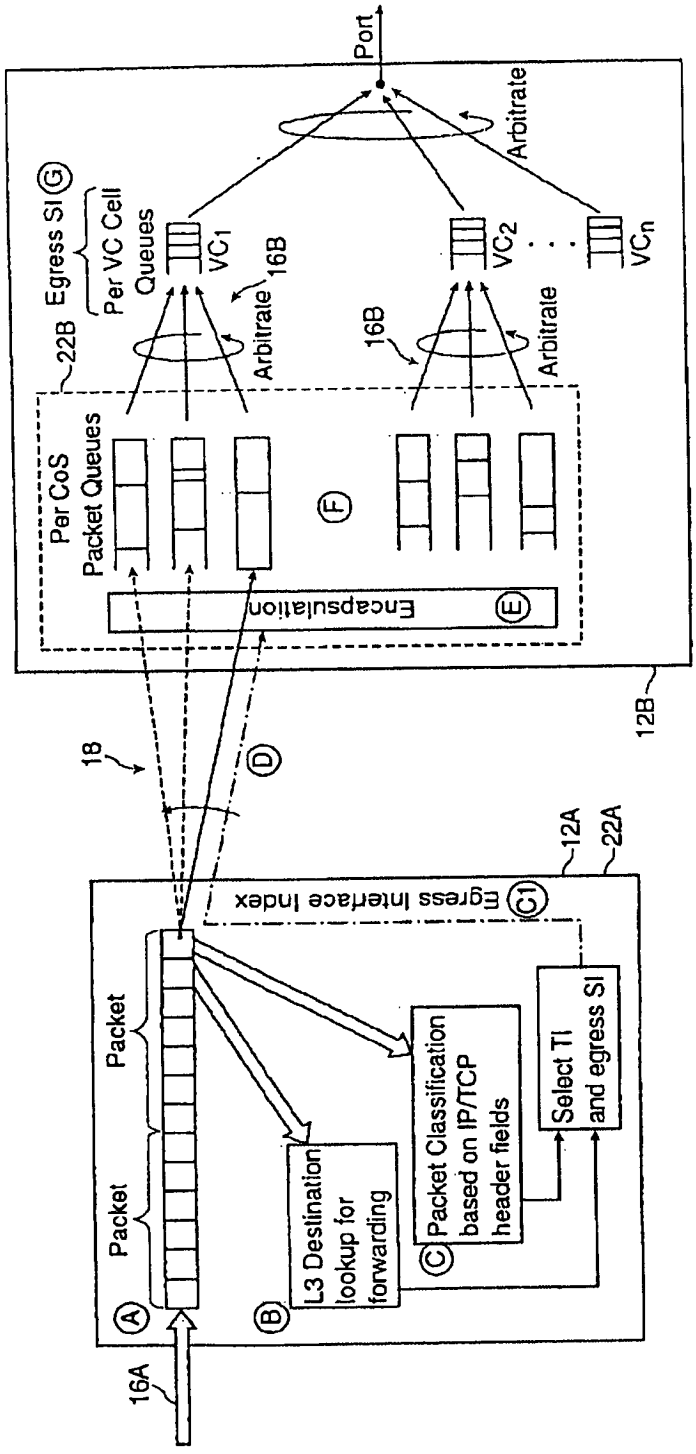

Figure 2

Fig. 3

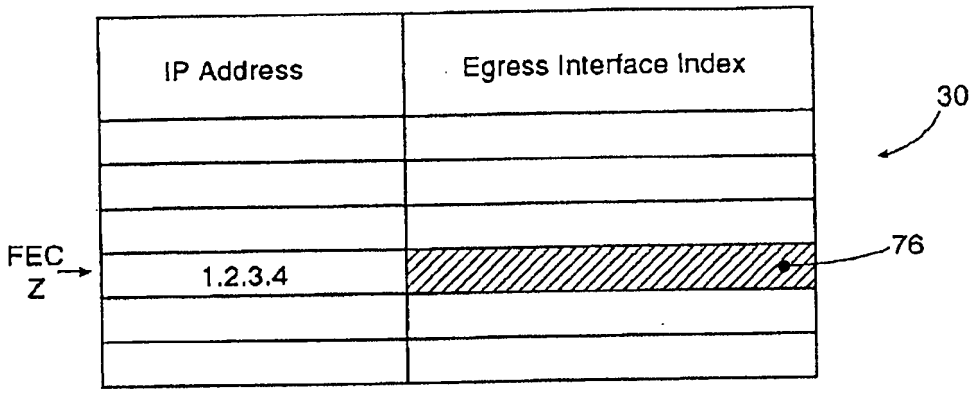

Figure 3

Fig. 4

Fig. 4

Parameter	Description	Default	NMTI	SNMP	NCI	CLI
ID number	A identification number assigned for this SI; unique within a subslot. This value is internally assigned and cannot be changed. This is 5 digit number field.	None	R	-	-	-
Endpoint	The ATM endpoint (shelf-slot-subslot-port; VPI/VCI) used by the SI.	None	R/W	-	-	-
Name	Name of the SI. This is 16 character text string field	Empty	R/W	-	-	-
Application	Application(s) provided by this SI. This is a boolean vector (i.e. bit map) indicating whether each of forwarding, routing and LDP is enabled.	Forward	R/W	-	-	-
Address type	Type of the IP address field. Valid type supported are unnumbered and IPv4.	Un-numbered	R/W	-	-	-
IP address	The IP address of the service interface. Represented to the user in standard "dotted decimal" format. "Illegal" IP addresses (e.g. 0.0.0.0, 255.255.255.255) are blocked.	Un-assigned	R/W	-	-	-
IP address prefix length	Number of bits in the IP address which constitute the (sub)network ID. A number in the range of 0..32.	None	R/W	-	-	-
Neighbour address type	Type of the neighbour IP address field. Valid type supported are unnumbered and IPv4.	Un-numbered	R/W	-	-	-
Neighbour IP address	The IP address used at the termination of the SI at the neighbouring router. Represented to the user in standard "dotted decimal" format. "Illegal" IP addresses (e.g. 0.0.0.0, 255.255.255.255) are blocked.	Un-assigned	R/W	-	-	-
Encapsulation	Encapsulation used on the SI. (RFC1483 LLC/SNAP routed IP, RFC1483 NULL)	RFC1483 NULL	R/W	-	-	-
MTU	Maximum Transmission Unit.	2016 octets	R	-	-	-
Ingress traffic contracts	An ingress traffic contract structure consists of an action (disable, tag, discard), a committed information rate (in b/s) and a burst size (in bytes). Eight ingress traffic contract structures are contained in each SI; each applies to a CoS.	disable CIR 0 BS 0	R/W	-	-	-
Status	Status of the service interface. (Up, Down).	Down	R	-	-	-

Service Interface Parameters

Fig. 5

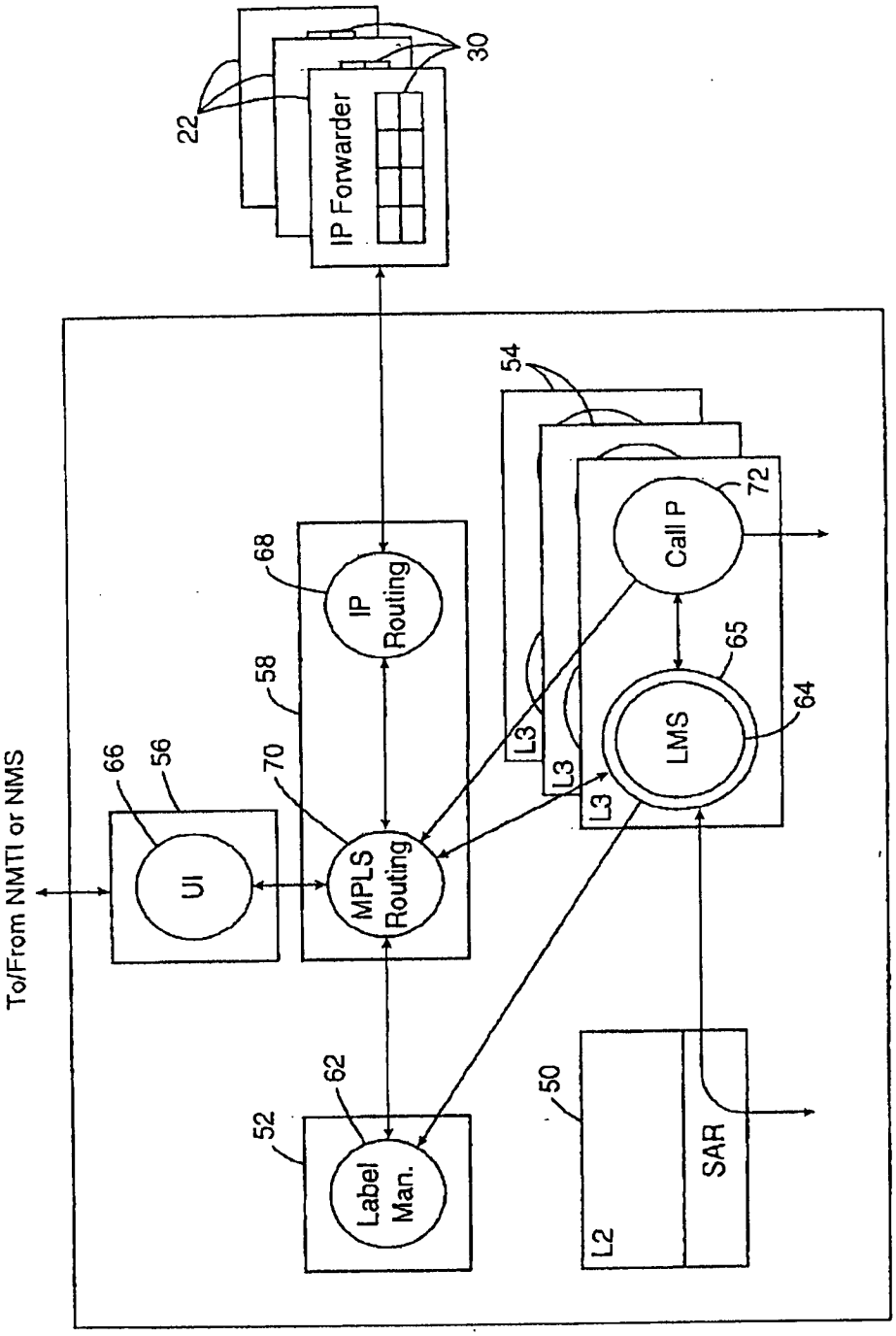

Figure 5

Fig. 6

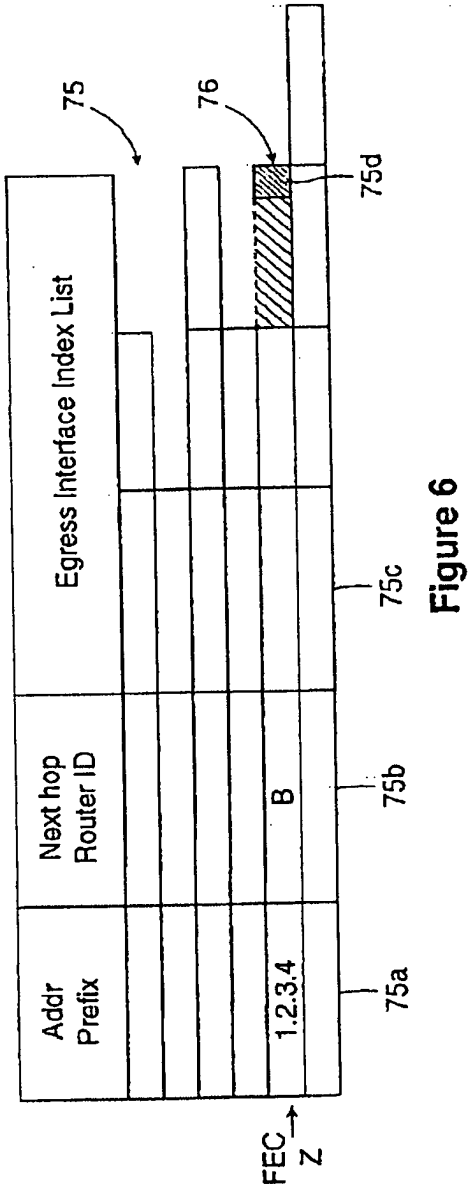

Figure 6

Fig. 7

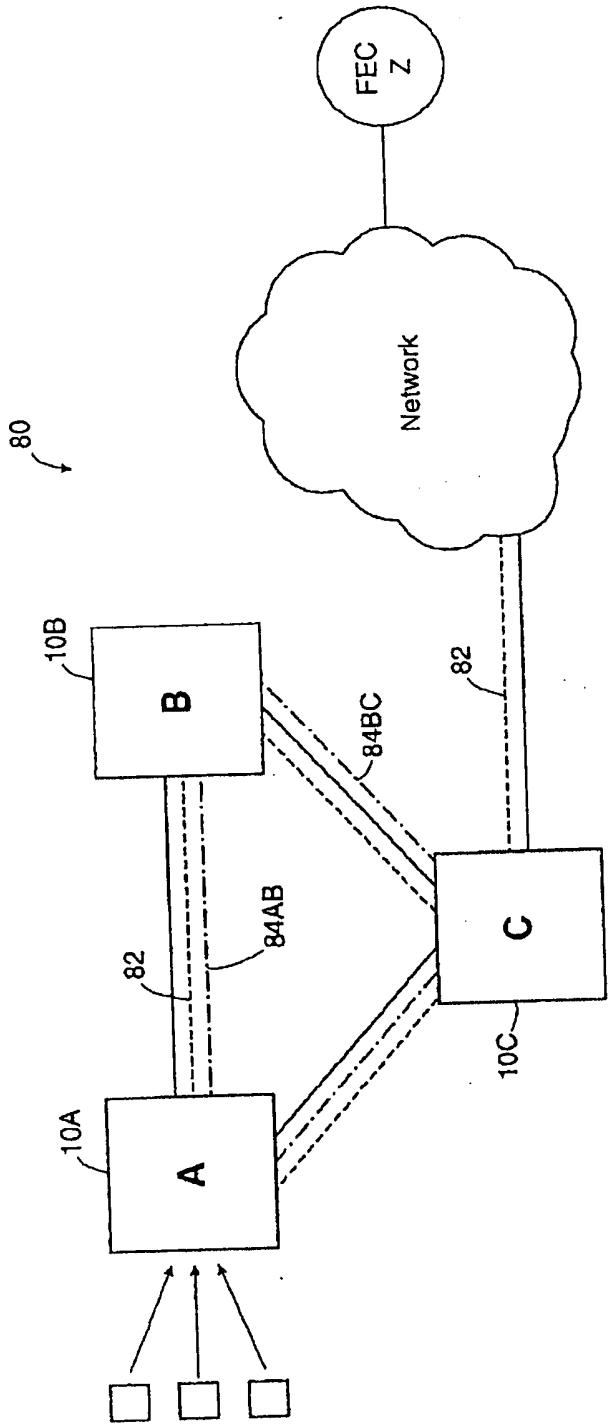

Figure 7

Fig. 8

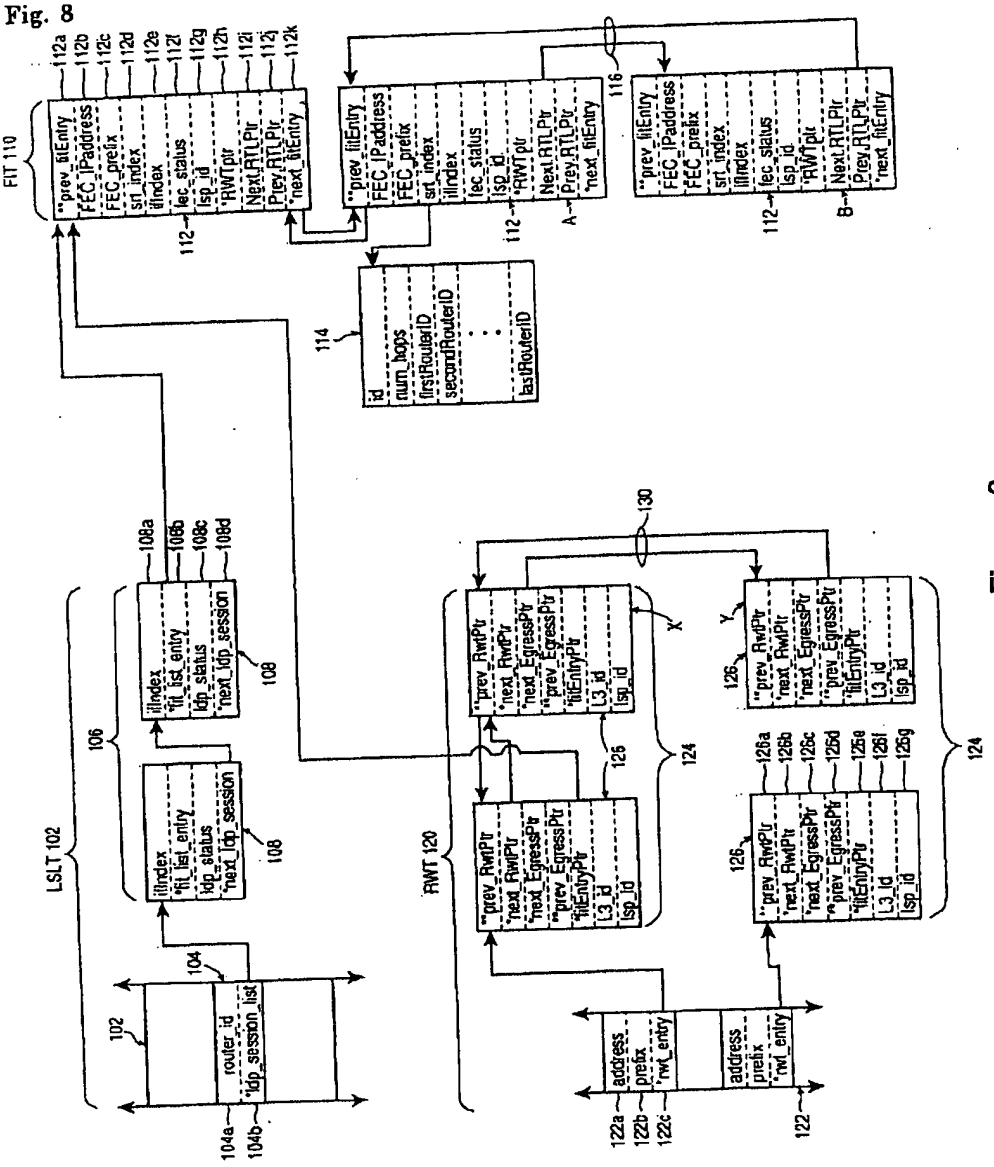

Figure 8

Fig. 8A

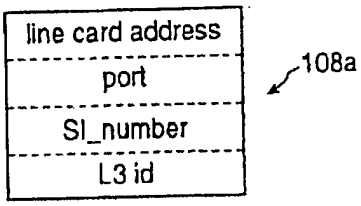

Figure 8A

Fig. 8B

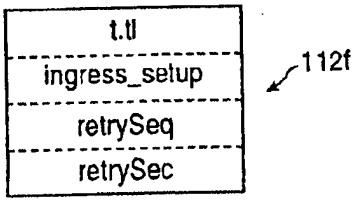

Figure 8B

Fig. 9

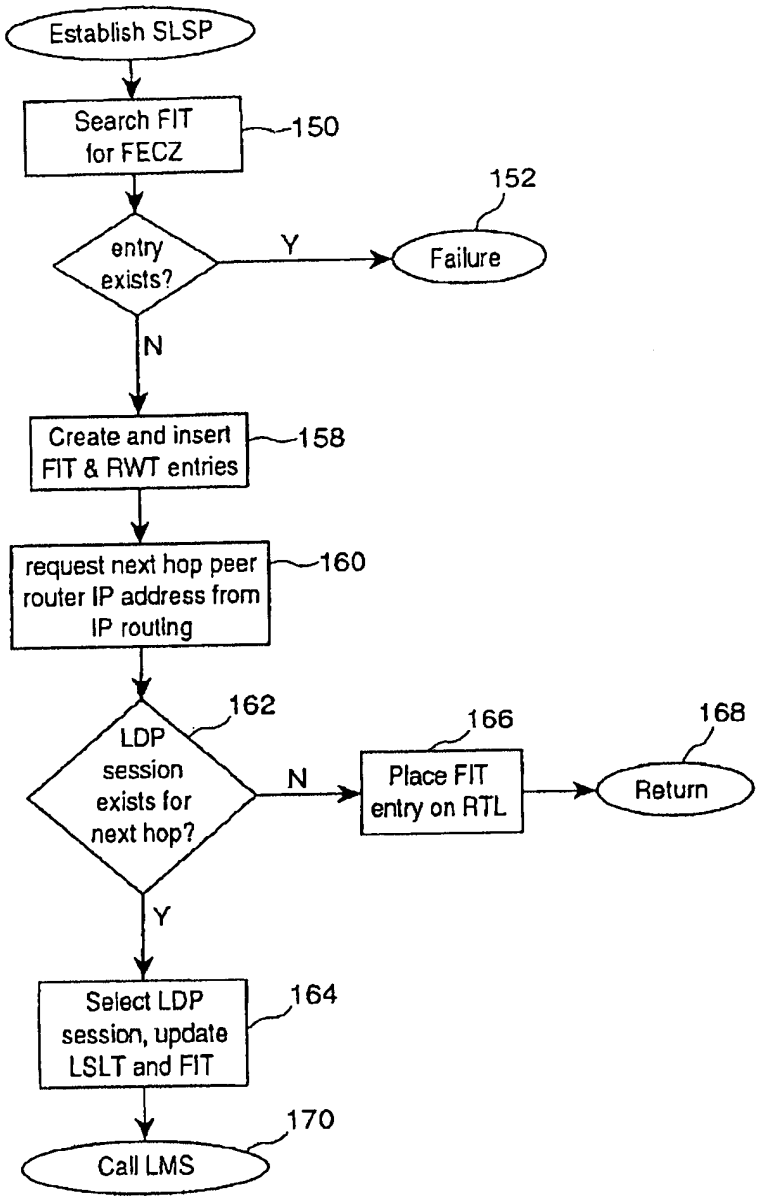

Figure 9

Fig. 10

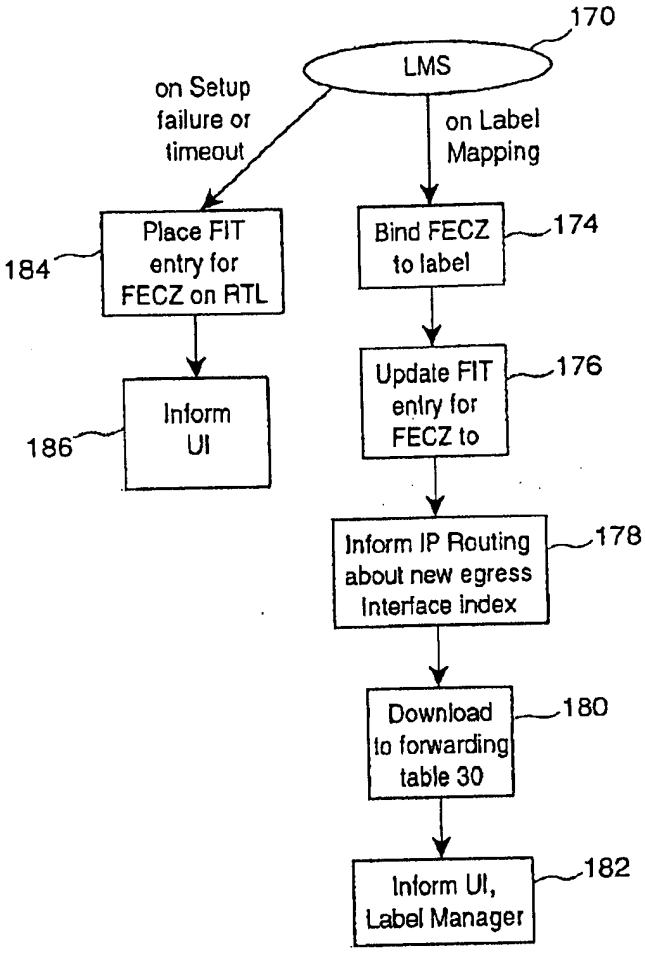

Figure 10

Fig. 11

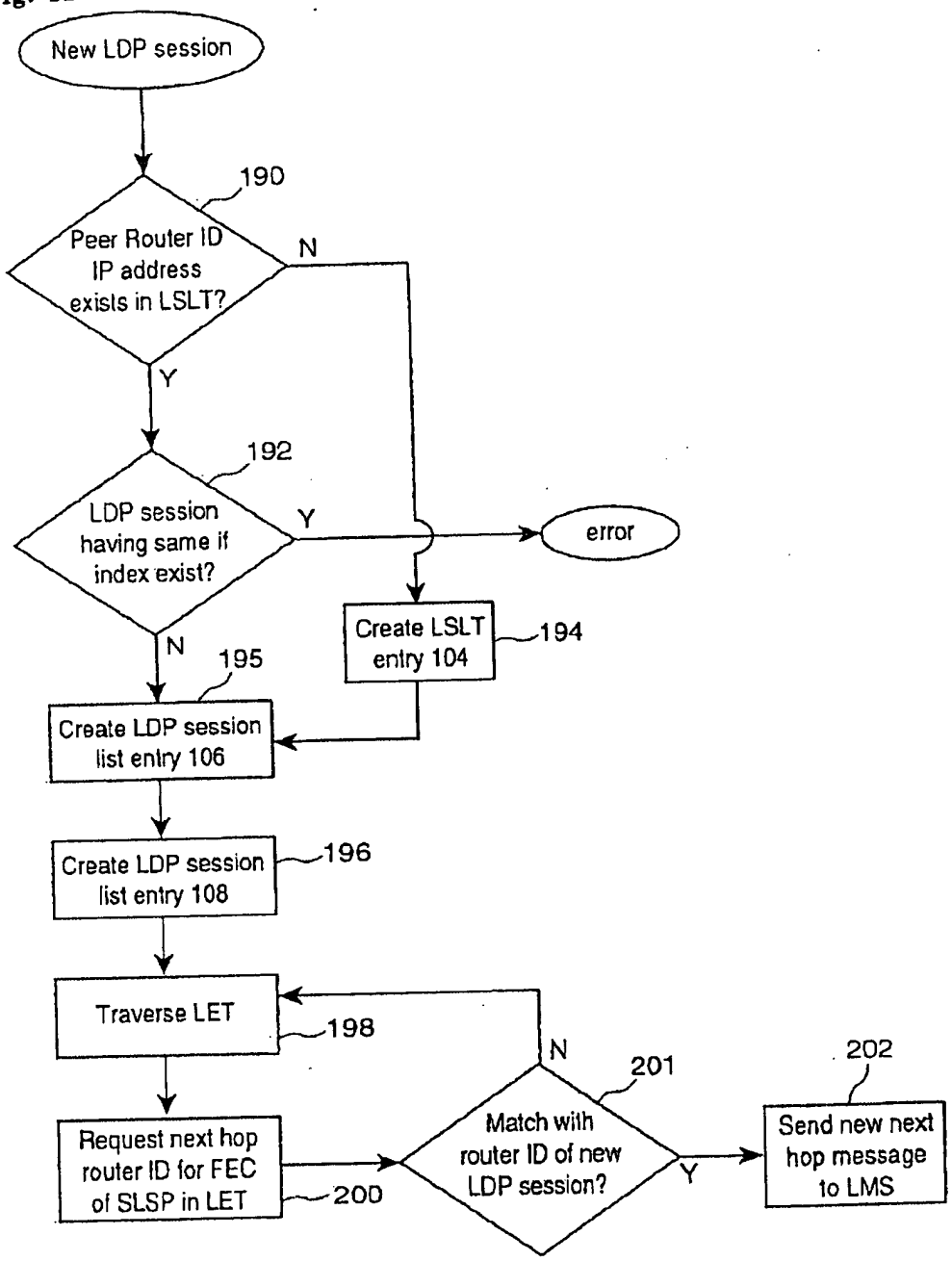

Figure 11

1. Abstract

A method of timing an attempt to establish a connection path between a first and second node in a communications network is provided. The method initiates the attempt to establish a connection path after a period of time has elapsed wherein the period of time is greater than another period of time which had previously elapsed between two previous attempts, if any, to establish the connection.

2. Representative Drawing

Fig. 1